GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 14:05:49 ; Search time 21 Seconds
(without alignments)
40.296 Million cell updates/sec

Title: US-09-651-685A-5
Perfect score: 114
Sequence: 1 CCYDGASVNNDETCEQRAAR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	109	95.6	73	4	US-09-097-094-3	Sequence 3, Appli
2	109	95.6	74	1	US-08-463-224-3	Sequence 3, Appli
3	109	95.6	74	2	US-08-463-377-3	Sequence 3, Appli
4	109	95.6	74	3	US-09-246-500B-10	Sequence 10, Appl
5	109	95.6	1676	4	US-08-487-283A-2	Sequence 2, Appli
6	94	82.5	74	3	US-09-246-500B-12	Sequence 12, Appl
7	93	81.6	74	3	US-09-246-500B-11	Sequence 11, Appl
8	88	77.2	77	3	US-09-246-500B-14	Sequence 14, Appl
9	84	73.7	77	3	US-09-246-500B-13	Sequence 13, Appl
10	55.5	48.7	76	3	US-09-246-500B-8	Sequence 8, Appli
11	51.5	45.2	77	3	US-09-246-500B-6	Sequence 6, Appli
12	50.5	44.3	77	3	US-09-246-500B-7	Sequence 7, Appli
13	48.5	42.5	76	3	US-09-246-500B-9	Sequence 9, Appli
14	44	38.6	237	4	US-09-328-352-8181	Sequence 8181, Ap
15	43.5	38.2	1651	1	US-08-447-411-2	Sequence 2, Appli
16	42	36.8	71	2	US-08-972-008-4	Sequence 4, Appli
17	42	36.8	71	4	US-09-267-409-4	Sequence 4, Appli
18	42	36.8	241	3	US-08-808-148-1	Sequence 1, Appli
19	42	36.8	241	3	US-09-020-956-114	Sequence 114, App
20	42	36.8	241	3	US-09-030-607-114	Sequence 114, App
21	42	36.8	241	4	US-09-439-313-114	Sequence 114, App
22	42	36.8	241	4	US-09-352-616A-114	Sequence 114, App
23	42	36.8	241	4	US-09-232-149A-114	Sequence 114, App
24	42	36.8	263	2	US-08-972-008-2	Sequence 2, Appli
25	42	36.8	263	4	US-09-141-027-2	Sequence 2, Appli
26	42	36.8	263	4	US-09-267-409-2	Sequence 2, Appli
27	42	36.8	263	4	US-09-617-804-2	Sequence 2, Appli
28	41.5	36.4	486	4	US-09-914-259-35	Sequence 35, Appl
29	41.5	36.4	1333	1	US-08-447-411-76	Sequence 76, Appl
30	41.5	36.4	1333	2	US-08-662-227-34	Sequence 34, Appl
31	41.5	36.4	1333	4	US-09-017-947-34	Sequence 34, Appl
32	41	36.0	363	2	US-08-483-101-16	Sequence 16, Appl
33	41	36.0	641	4	US-09-687-538B-8	Sequence 8, Appli
34	40	35.1	115	4	US-09-107-532A-6191	Sequence 6191, Ap
35	40	35.1	160	4	US-09-252-991A-22711	Sequence 22711, A
36	40	35.1	289	4	US-09-107-532A-4580	Sequence 4580, Ap
37	40	35.1	751	2	US-08-836-443-3	Sequence 3, Appli
38	40	35.1	802	4	US-09-632-098-2	Sequence 2, Appli
39	40	35.1	812	4	US-09-632-098-4	Sequence 4, Appli
40	40	35.1	1940	2	US-08-644-271-30	Sequence 30, Appl
41	40	35.1	1940	4	US-09-077-955-34	Sequence 34, Appl
42	40	35.1	4544	1	US-08-469-486-52	Sequence 52, Appl
43	40	35.1	4544	2	US-08-469-658-52	Sequence 52, Appl
44	39.5	34.6	256	4	US-09-325-932A-57	Sequence 57, Appl
45	39	34.2	34	2	US-08-867-087B-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1
US-09-097-094-3
; Sequence 3, Application US/09097094
; Patent No. 6326468
; GENERAL INFORMATION:
; APPLICANT: Canne, Lynne
; APPLICANT: Kent, Stephen B.H.
; APPLICANT: Simon, Reyna
; TITLE OF INVENTION: Solid Phase Native Chenical Ligation of Unprotected or
; TITLE OF INVENTION: N-Terminal Cysteine Protected Peptides in Aqueous
; TITLE OF INVENTION: Solution
; FILE REFERENCE: GRFN-023/01US
; CURRENT APPLICATION NUMBER: US/09/097,094
; CURRENT FILING DATE: 1998-06-12
; EARLIER APPLICATION NUMBER: 60/049,553
; EARLIER FILING DATE: 1997-06-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-097-094-3

Query Match 95.6%; Score 109; DB 4; Length 73;
Best Local Similarity 95.0%; Pred. No. 8e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||||||||||||
Db         20 CCYDGACVNNDETCEQRAAR 39
```

RESULT 2
US-08-463-224-3
; Sequence 3, Application US/08463224
; Patent No. 5807824
; GENERAL INFORMATION:
; APPLICANT: van Oostrum, Jan
; APPLICANT: Boyar, William C.
; APPLICANT: Galakatos, Nicholas G.
; APPLICANT: Schmitz, Albert
; APPLICANT: van Heeke, Gino
; TITLE OF INVENTION: C5a Receptor Antagonists Having
; TITLE OF INVENTION: Substantially No. 5807824Agonist Activity
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik
; STREET: 600 South Avenue West
; CITY: Westfield

```
;       STATE:  NJ
;       COUNTRY:  USA
;       ZIP:  07090
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/463,224
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Foley, Shawn P.
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  908-654-5000
;       TELEFAX:  908-654-7866
;       TELEX:  139-125
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  74 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-463-224-3

  Query Match             95.6%;  Score 109;  DB 1;  Length 74;
  Best Local Similarity   95.0%;  Pred. No. 8.1e-10;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||||||||||||
Db         21 CCYDGACVNNDETCEQRAAR 40


RESULT 3
US-08-463-377-3
; Sequence 3, Application US/08463377
; Patent No. 5837499
;  GENERAL INFORMATION:
;     APPLICANT:  van Oostrum, Jan
;     APPLICANT:  Boyar, William C.
;     APPLICANT:  Galakatos, Nicholas G.
;     APPLICANT:  Schmitz, Albert
;     APPLICANT:  van Heeke, Gino
;     TITLE OF INVENTION:  C5a Receptor Antagonists Having
;     TITLE OF INVENTION:  Substantially No. 5837499Agonist Activity
;     NUMBER OF SEQUENCES:  67
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Lerner, David, Littenberg, Krumholz & Mentlik
;       STREET:  600 South Avenue West
;       CITY:  Westfield
;       STATE:  NJ
;       COUNTRY:  USA
;       ZIP:  07090
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/463,377
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Foley, Shawn P.
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  908-654-5000
;       TELEFAX:  908-654-7866
;       TELEX:  139-125
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  74 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-463-377-3

  Query Match             95.6%;  Score 109;  DB 2;  Length 74;
  Best Local Similarity   95.0%;  Pred. No. 8.1e-10;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||||||||||||
Db         21 CCYDGACVNNDETCEQRAAR 40


RESULT 4
US-09-246-500B-10
; Sequence 10, Application US/09246500B
; Patent No. 6235494
; GENERAL INFORMATION:
;  APPLICANT:  Hugli, Tony E.
;  TITLE OF INVENTION:  Substrates for Assessing Mannan-Binding
;  TITLE OF INVENTION:  Protein-Associated Serine Protease Activity and Methods
;  TITLE OF INVENTION:  Using the Substrates
;  FILE REFERENCE:  24730-2204
;  CURRENT APPLICATION NUMBER: US/09/246,500B
;  CURRENT FILING DATE:  1999-02-08
;  NUMBER OF SEQ ID NOS:  19
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 74
;   TYPE: PRT
;   ORGANISM: Human C5a Anaphylatoxin
;   FEATURE:
;   NAME/KEY: CARBOHYD
;   LOCATION: (64)...(66)
US-09-246-500B-10

  Query Match             95.6%;  Score 109;  DB 3;  Length 74;
  Best Local Similarity   95.0%;  Pred. No. 8.1e-10;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||||||||||||
Db         21 CCYDGACVNNDETCEQRAAR 40


RESULT 5
US-08-487-283A-2
; Sequence 2, Application US/08487283A
; Patent No. 6355245
;  GENERAL INFORMATION:
;    APPLICANT:  Evans, Mark J.
;    APPLICANT:  Matis, Louis A.
;    APPLICANT:  Mueller, Eileen Elliott
;    APPLICANT:  Nye, Steven H.
;    APPLICANT:  Rollins, Scott
;    APPLICANT:  Rother, Russell P.
;    APPLICANT:  Springhorn, Jeremy P.
;    APPLICANT:  Squinto, Stephen P.
;    APPLICANT:  Thomas, Thomas C.
;    APPLICANT:  Wilkins, James A.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR THE TREATMENT
;    TITLE OF INVENTION:  OF INFLAMMATORY DISEASES
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seth A. Fidel
;      STREET:  25 Science Park (Alexion)
;      CITY:  New Haven
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06511
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.4Mb storage
```

```
;      COMPUTER:  Macintosh Cetris 610
;      OPERATING SYSTEM:  System 7
;      SOFTWARE:  WordPerfect 3.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,283A
;      FILING DATE:  June 7, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/236,208
;      FILING DATE:  02-MAY-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seth A. Fidel.
;      REGISTRATION NUMBER:  38,449
;      REFERENCE/DOCKET NUMBER:  ALX-152.1 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203)776-1790
;      TELEFAX:  (203)772-3655
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1676 Amino Acids
;      TYPE:  Amino Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:
;      DESCRIPTION:  Pro-C5 Polytpeptide
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    PUBLICATION INFORMATION:
;      AUTHORS:  Haviland, D.L.
;      AUTHORS:  Haviland, J.C.
;      AUTHORS:  Fleischer, D.T.
;      AUTHORS:  Hunt, A.
;      AUTHORS:  Wetsel, R.A.
;      TITLE:  Complete cDNA Sequence of Human
; Patent No. 6355245
;      TITLE:  Complement Pro-C5
;      JOURNAL:  Journal of Immunology
;      VOLUME:  146
;      PAGES:  362-368
;      DATE:  1991
US-08-487-283A-2

  Query Match             95.6%;  Score 109;  DB 4;  Length 1676;
  Best Local Similarity   95.0%;  Pred. No. 2.3e-08;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||||||||||||||
Db        698 CCYDGACVNNDETCEQRAAR 717


RESULT 6
US-09-246-500B-12
; Sequence 12, Application US/09246500B
; Patent No. 6235494
; GENERAL INFORMATION:
;  APPLICANT:  Hugli, Tony E.
;  TITLE OF INVENTION:  Substrates for Assessing Mannan-Binding
;  TITLE OF INVENTION:  Protein-Associated Serine Protease Activity and Methods
;  TITLE OF INVENTION:  Using the Substrates
;  FILE REFERENCE:  24730-2204
;  CURRENT APPLICATION NUMBER: US/09/246,500B
;  CURRENT FILING DATE:  1999-02-08
;  NUMBER OF SEQ ID NOS:  19
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 74
;   TYPE: PRT
;   ORGANISM: Cow C5a Anaphylatoxin
US-09-246-500B-12

  Query Match             82.5%;  Score 94;  DB 3;  Length 74;
  Best Local Similarity   80.0%;  Pred. No. 1.6e-07;
  Matches   16;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| |:|||||:||||
Db         21 CCYDGAHRNDDETCEERAAR 40


RESULT 7
US-09-246-500B-11
; Sequence 11, Application US/09246500B
; Patent No. 6235494
; GENERAL INFORMATION:
;  APPLICANT:  Hugli, Tony E.
;  TITLE OF INVENTION:  Substrates for Assessing Mannan-Binding
;  TITLE OF INVENTION:  Protein-Associated Serine Protease Activity and Methods
;  TITLE OF INVENTION:  Using the Substrates
;  FILE REFERENCE:  24730-2204
;  CURRENT APPLICATION NUMBER: US/09/246,500B
;  CURRENT FILING DATE:  1999-02-08
;  NUMBER OF SEQ ID NOS:  19
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 74
;   TYPE: PRT
;   ORGANISM: Pig C5a Anaphylatoxin
;   FEATURE:
;   NAME/KEY: CARBOHYD
;   LOCATION: (64)...(66)
US-09-246-500B-11

  Query Match             81.6%;  Score 93;  DB 3;  Length 74;
  Best Local Similarity   80.0%;  Pred. No. 2.3e-07;
  Matches   16;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| |:|||||:||||
Db         21 CCYDGAYRNDDETCEERAAR 40


RESULT 8
US-09-246-500B-14
; Sequence 14, Application US/09246500B
; Patent No. 6235494
; GENERAL INFORMATION:
;  APPLICANT:  Hugli, Tony E.
;  TITLE OF INVENTION:  Substrates for Assessing Mannan-Binding
;  TITLE OF INVENTION:  Protein-Associated Serine Protease Activity and Methods
;  TITLE OF INVENTION:  Using the Substrates
;  FILE REFERENCE:  24730-2204
;  CURRENT APPLICATION NUMBER: US/09/246,500B
;  CURRENT FILING DATE:  1999-02-08
;  NUMBER OF SEQ ID NOS:  19
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 77
;   TYPE: PRT
;   ORGANISM: Mouse C5a Anaphylatoxin
US-09-246-500B-14

  Query Match             77.2%;  Score 88;  DB 3;  Length 77;
  Best Local Similarity   80.0%;  Pred. No. 1.4e-06;
  Matches   16;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||  ||||:||||
Db         24 CCYDGARVNFYETCEERAAR 43


RESULT 9
US-09-246-500B-13
; Sequence 13, Application US/09246500B
; Patent No. 6235494
```

```
; GENERAL INFORMATION:
;  APPLICANT:  Hugli, Tony E.
;  TITLE OF INVENTION:  Substrates for Assessing Mannan-Binding
;  TITLE OF INVENTION:  Protein-Associated Serine Protease Activity and Methods
;  TITLE OF INVENTION:  Using the Substrates
;  FILE REFERENCE:  24730-2204
;  CURRENT APPLICATION NUMBER: US/09/246,500B
;  CURRENT FILING DATE:  1999-02-08
;  NUMBER OF SEQ ID NOS:  19
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 77
;   TYPE: PRT
;   ORGANISM: Rat C5a Anaphylatoxin
;   FEATURE:
;   NAME/KEY: CARBOHYD
;   LOCATION: (67)...(69)
US-09-246-500B-13

  Query Match             73.7%;  Score 84;  DB 3;  Length 77;
  Best Local Similarity   75.0%;  Pred. No. 5.6e-06;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||  |  |||||| ||
Db         24 CCYDGARENKYETCEQRVAR 43


RESULT 10
US-09-246-500B-8
; Sequence 8, Application US/09246500B
; Patent No. 6235494
; GENERAL INFORMATION:
;  APPLICANT:  Hugli, Tony E.
;  TITLE OF INVENTION:  Substrates for Assessing Mannan-Binding
;  TITLE OF INVENTION:  Protein-Associated Serine Protease Activity and Methods
;  TITLE OF INVENTION:  Using the Substrates
;  FILE REFERENCE:  24730-2204
;  CURRENT APPLICATION NUMBER: US/09/246,500B
;  CURRENT FILING DATE:  1999-02-08
;  NUMBER OF SEQ ID NOS:  19
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 76
;   TYPE: PRT
;   ORGANISM: Rat C4a Anaphylatoxin
;   FEATURE:
;   NAME/KEY: CARBOHYD
;   LOCATION: (66)...(68)
US-09-246-500B-8

  Query Match             48.7%;  Score 55.5;  DB 3;  Length 76;
  Best Local Similarity   57.1%;  Pred. No. 0.13;
  Matches   12;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          1 CCYDG-ASVNNDETCEQRAAR 20
              || ||   :   ||||||||
Db         23 CCQDGMTKLPMARTCEQRAAR 43


RESULT 11
US-09-246-500B-6
; Sequence 6, Application US/09246500B
; Patent No. 6235494
; GENERAL INFORMATION:
;  APPLICANT:  Hugli, Tony E.
;  TITLE OF INVENTION:  Substrates for Assessing Mannan-Binding
;  TITLE OF INVENTION:  Protein-Associated Serine Protease Activity and Methods
;  TITLE OF INVENTION:  Using the Substrates
;  FILE REFERENCE:  24730-2204
;  CURRENT APPLICATION NUMBER: US/09/246,500B
;  CURRENT FILING DATE:  1999-02-08
;  NUMBER OF SEQ ID NOS:  19
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 77
;   TYPE: PRT
;   ORGANISM: Human C4a Anaphylatoxin
US-09-246-500B-6

  Query Match             45.2%;  Score 51.5;  DB 3;  Length 77;
  Best Local Similarity   52.4%;  Pred. No. 0.53;
  Matches   11;  Conservative    3;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 CCYDGAS-VNNDETCEQRAAR 20
              || ||  : :    :|||||||
Db         23 CCQDGVTRLPMMRSCEQRAAR 43


RESULT 12
US-09-246-500B-7
; Sequence 7, Application US/09246500B
; Patent No. 6235494
; GENERAL INFORMATION:
;  APPLICANT:  Hugli, Tony E.
;  TITLE OF INVENTION:  Substrates for Assessing Mannan-Binding
;  TITLE OF INVENTION:  Protein-Associated Serine Protease Activity and Methods
;  TITLE OF INVENTION:  Using the Substrates
;  FILE REFERENCE:  24730-2204
;  CURRENT APPLICATION NUMBER: US/09/246,500B
;  CURRENT FILING DATE:  1999-02-08
;  NUMBER OF SEQ ID NOS:  19
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 77
;   TYPE: PRT
;   ORGANISM: Cow C4a Anaphylatoxin
US-09-246-500B-7

  Query Match             44.3%;  Score 50.5;  DB 3;  Length 77;
  Best Local Similarity   52.4%;  Pred. No. 0.75;
  Matches   11;  Conservative    3;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 CCYDGAS-VNNDETCEQRAAR 20
              || ||  : :    :|||||||
Db         23 CCQDGLTRLPMMRSCEQRAAR 43


RESULT 13
US-09-246-500B-9
; Sequence 9, Application US/09246500B
; Patent No. 6235494
; GENERAL INFORMATION:
;  APPLICANT:  Hugli, Tony E.
;  TITLE OF INVENTION:  Substrates for Assessing Mannan-Binding
;  TITLE OF INVENTION:  Protein-Associated Serine Protease Activity and Methods
;  TITLE OF INVENTION:  Using the Substrates
;  FILE REFERENCE:  24730-2204
;  CURRENT APPLICATION NUMBER: US/09/246,500B
;  CURRENT FILING DATE:  1999-02-08
;  NUMBER OF SEQ ID NOS:  19
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 76
;   TYPE: PRT
;   ORGANISM: Mouse C4a Anaphylatoxin
;   FEATURE:
;   NAME/KEY: CARBOHYD
;   LOCATION: (66)...(68)
US-09-246-500B-9

  Query Match             42.5%;  Score 48.5;  DB 3;  Length 76;
  Best Local Similarity   50.0%;  Pred. No. 1.5;
  Matches   10;  Conservative    3;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 CCYDGAS-VNNDETCEQRAA 19
```

```
              || || : :    :||||||
Db          23 CCQDGVTRLPMKRSCEQRAA 42


RESULT 14
US-09-328-352-8181
; Sequence 8181, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8181
;   LENGTH: 237
;   TYPE: PRT
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-8181

  Query Match             38.6%;  Score 44;  DB 4;  Length 237;
  Best Local Similarity   66.7%;  Pred. No. 25;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 DGASVNNDETCE 15
              | |:| |||||:
Db        221 DDAAVCNDETCD 232


RESULT 15
US-08-447-411-2
; Sequence 2, Application US/08447411
; Patent No. 5773243
;  GENERAL INFORMATION:
;    APPLICANT:  FRITZINGER, DAVID C.
;    APPLICANT:  BREDEHORST, REINHARD
;    APPLICANT:  VOGEL, CARL-WILHELM
;    TITLE OF INVENTION:  DNA ENCODING COBRA C3, CVF1, AND CVF2
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,411
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/043,747
;      FILING DATE:  07-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5773243man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-101-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1651 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-447-411-2

  Query Match             38.2%;  Score 43.5;  DB 1;  Length 1651;
  Best Local Similarity   52.6%;  Pred. No. 2.3e+02;
  Matches   10;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          1 CCYDGASVN-NDETCEQRA 18
              || ||   |   |||:||
Db        683 CCEDGMHENPMGYTCEKRA 701


Search completed: December  9, 2003, 14:08:59
Job time : 22 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 14:07:54 ; Search time 30 Seconds
(without alignments)
123.989 Million cell updates/sec

Title: US-09-651-685A-5
Perfect score: 114
Sequence: 1 CCYDGASVNNDETCEQRAAR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	114	100.0	20	10	US-09-878-603-5	Sequence 5, Appli
2	114	100.0	74	10	US-09-878-603-3	Sequence 3, Appli
3	111	97.4	20	10	US-09-878-603-68	Sequence 68, Appl
4	110	96.5	20	10	US-09-878-603-67	Sequence 67, Appl
5	110	96.5	20	10	US-09-878-603-69	Sequence 69, Appl
6	110	96.5	20	10	US-09-878-603-70	Sequence 70, Appl
7	110	96.5	20	10	US-09-878-603-71	Sequence 71, Appl
8	110	96.5	20	10	US-09-878-603-73	Sequence 73, Appl
9	109	95.6	19	10	US-09-878-603-54	Sequence 54, Appl
10	109	95.6	73	10	US-09-987-655-3	Sequence 3, Appli
11	109	95.6	73	10	US-09-987-675-3	Sequence 3, Appli
12	109	95.6	1251	9	US-09-778-927A-58	Sequence 58, Appl
13	109	95.6	1602	9	US-09-778-927A-59	Sequence 59, Appl
14	108	94.7	20	10	US-09-878-603-64	Sequence 64, Appl
15	108	94.7	20	10	US-09-878-603-65	Sequence 65, Appl
16	106	93.0	20	10	US-09-878-603-72	Sequence 72, Appl
17	105	92.1	18	10	US-09-878-603-55	Sequence 55, Appl
18	105	92.1	19	10	US-09-878-603-59	Sequence 59, Appl
19	102	89.5	20	10	US-09-878-603-66	Sequence 66, Appl
20	101	88.6	17	10	US-09-878-603-56	Sequence 56, Appl
21	98	86.0	20	10	US-09-878-603-74	Sequence 74, Appl
22	96	84.2	16	10	US-09-878-603-57	Sequence 57, Appl
23	96	84.2	18	10	US-09-878-603-60	Sequence 60, Appl
24	94	82.5	74	10	US-09-878-603-7	Sequence 7, Appli
25	93	81.6	74	10	US-09-878-603-8	Sequence 8, Appli
26	91	79.8	15	10	US-09-878-603-58	Sequence 58, Appl
27	89	78.1	17	10	US-09-878-603-61	Sequence 61, Appl
28	85	74.6	354	14	US-10-039-050-2	Sequence 2, Appli
29	84	73.7	77	10	US-09-878-603-1	Sequence 1, Appli
30	83	72.8	16	10	US-09-878-603-62	Sequence 62, Appl
31	77	67.5	15	10	US-09-878-603-63	Sequence 63, Appl
32	51	44.7	20	10	US-09-878-603-2	Sequence 2, Appli
33	45	39.5	141	12	US-10-238-075-831	Sequence 831, App
34	44	38.6	715	12	US-10-021-660-110	Sequence 110, App
35	44	38.6	884	12	US-10-032-585-7212	Sequence 7212, Ap
36	43.5	38.2	241	15	US-10-097-340-326	Sequence 326, App
37	43.5	38.2	241	15	US-10-205-823-425	Sequence 425, App
38	42	36.8	233	10	US-09-981-876-137	Sequence 137, App
39	42	36.8	233	11	US-09-148-545-137	Sequence 137, App
40	42	36.8	241	9	US-09-759-143-114	Sequence 114, App
41	42	36.8	241	9	US-09-780-669-114	Sequence 114, App
42	42	36.8	241	9	US-09-030-606-114	Sequence 114, App
43	42	36.8	241	9	US-09-822-827-114	Sequence 114, App
44	42	36.8	241	9	US-09-115-453-114	Sequence 114, App
45	42	36.8	241	10	US-09-232-880-114	Sequence 114, App

ALIGNMENTS

RESULT 1
US-09-878-603-5
; Sequence 5, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-878-603-5

Query Match 100.0%; Score 114; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db          1 CCYDGASVNNDETCEQRAAR 20
```

RESULT 2
US-09-878-603-3
; Sequence 3, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.

```
;  APPLICANT:  Huber-Lang, Markus
;  APPLICANT:  Sarma, Vidya
;  APPLICANT:  Czermak, Boris
;  TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
;  FILE REFERENCE: UM-03783
;  CURRENT APPLICATION NUMBER: US/09/878,603
;  CURRENT FILING DATE:  2001-06-11
;  PRIOR APPLICATION NUMBER: 09/387,671
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 74
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-878-603-3

  Query Match             100.0%;  Score 114;  DB 10;  Length 74;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db         21 CCYDGASVNNDETCEQRAAR 40


RESULT 3
US-09-878-603-68
; Sequence 68, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
;  APPLICANT: Ward, Peter A.
;  APPLICANT:  Huber-Lang, Markus
;  APPLICANT:  Sarma, Vidya
;  APPLICANT:  Czermak, Boris
;  TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
;  FILE REFERENCE: UM-03783
;  CURRENT APPLICATION NUMBER: US/09/878,603
;  CURRENT FILING DATE:  2001-06-11
;  PRIOR APPLICATION NUMBER: 09/387,671
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
US-09-878-603-68

  Query Match             97.4%;  Score 111;  DB 10;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 2.9e-10;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||:|||||||||||||
Db          1 CCYDGATVNNDETCEQRAAR 20


RESULT 4
US-09-878-603-67
; Sequence 67, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
;  APPLICANT: Ward, Peter A.
;  APPLICANT:  Huber-Lang, Markus
;  APPLICANT:  Sarma, Vidya
;  APPLICANT:  Czermak, Boris
;  TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
;  FILE REFERENCE: UM-03783
;  CURRENT APPLICATION NUMBER: US/09/878,603
;  CURRENT FILING DATE:  2001-06-11
;  PRIOR APPLICATION NUMBER: 09/387,671
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 67
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
US-09-878-603-67

  Query Match             96.5%;  Score 110;  DB 10;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 4.2e-10;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||:|||||||
Db          1 CCYDGASVNNDESCEQRAAR 20


RESULT 5
US-09-878-603-69
; Sequence 69, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
;  APPLICANT: Ward, Peter A.
;  APPLICANT:  Huber-Lang, Markus
;  APPLICANT:  Sarma, Vidya
;  APPLICANT:  Czermak, Boris
;  TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
;  FILE REFERENCE: UM-03783
;  CURRENT APPLICATION NUMBER: US/09/878,603
;  CURRENT FILING DATE:  2001-06-11
;  PRIOR APPLICATION NUMBER: 09/387,671
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
US-09-878-603-69

  Query Match             96.5%;  Score 110;  DB 10;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 4.2e-10;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||| |||||||||||||||
Db          1 CCYDGVSVNNDETCEQRAAR 20


RESULT 6
US-09-878-603-70
; Sequence 70, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
;  APPLICANT: Ward, Peter A.
;  APPLICANT:  Huber-Lang, Markus
;  APPLICANT:  Sarma, Vidya
;  APPLICANT:  Czermak, Boris
;  TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
;  FILE REFERENCE: UM-03783
;  CURRENT APPLICATION NUMBER: US/09/878,603
;  CURRENT FILING DATE:  2001-06-11
;  PRIOR APPLICATION NUMBER: 09/387,671
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
```

; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-70

Query Match 96.5%; Score 110; DB 10; Length 20;
Best Local Similarity 95.0%; Pred. No. 4.2e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||| ||||||||||||
Db 1 CCYDGASANNDETCEQRAAR 20

RESULT 7
US-09-878-603-71
; Sequence 71, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-71

Query Match 96.5%; Score 110; DB 10; Length 20;
Best Local Similarity 95.0%; Pred. No. 4.2e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
|||||||||||||||| ||
Db 1 CCYDGASVNNDETCEQRVAR 20

RESULT 8
US-09-878-603-73
; Sequence 73, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-73

Query Match 96.5%; Score 110; DB 10; Length 20;
Best Local Similarity 95.0%; Pred. No. 4.2e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
|||||||||||||||| |
Db 1 CCYDGASVNNDETCEQRAVR 20

RESULT 9
US-09-878-603-54
; Sequence 54, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-54

Query Match 95.6%; Score 109; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 5.6e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAA 19
|||||||||||||||||||
Db 1 CCYDGASVNNDETCEQRAA 19

RESULT 10
US-09-987-655-3
; Sequence 3, Application US/09987655
; Patent No. US20020132975A1
; GENERAL INFORMATION:
; APPLICANT: Canne, Lynne
; APPLICANT: Kent, Stephen B.H.
; APPLICANT: Simon, Reyna
; TITLE OF INVENTION: Solid Phase Native Chemical Ligation of Unprotected or
; TITLE OF INVENTION: N-Terminal Cysteine Protected Peptides in Aqueous
; TITLE OF INVENTION: Solution
; FILE REFERENCE: GRFN-023/01US
; CURRENT APPLICATION NUMBER: US/09/987,655
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 09/097,094
; PRIOR FILING DATE: 1998-06-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-655-3

Query Match 95.6%; Score 109; DB 10; Length 73;
Best Local Similarity 95.0%; Pred. No. 2.4e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||||||||||||||
Db         20 CCYDGACVNNDETCEQRAAR 39
```

RESULT 11
US-09-987-675-3
; Sequence 3, Application US/09987675
; Patent No. US20020169282A1
; GENERAL INFORMATION:
; APPLICANT: Canne, Lynne
; APPLICANT: Kent, Stephen B.H.
; APPLICANT: Simon, Reyna
; TITLE OF INVENTION: Solid Phase Native Chemical Ligation of Unprotected or
; TITLE OF INVENTION: N-Terminal Cysteine Protected Peptides in Aqueous
; TITLE OF INVENTION: Solution
; FILE REFERENCE: GRFN-023/01US
; CURRENT APPLICATION NUMBER: US/09/987,675
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 09/097,094
; PRIOR FILING DATE: 1998-06-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-675-3

Query Match 95.6%; Score 109; DB 10; Length 73;
Best Local Similarity 95.0%; Pred. No. 2.4e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||||||||||||||
Db         20 CCYDGACVNNDETCEQRAAR 39
```

RESULT 12
US-09-778-927A-58
; Sequence 58, Application US/09778927A
; Patent No. US20020068342A1
; GENERAL INFORMATION:
; APPLICANT: KHOSRAVI, Rami et al.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL
; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING
; FILE REFERENCE: 2786-0160P
; CURRENT APPLICATION NUMBER: US/09/778,927A
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: IL 134453
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: IL135341
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 1251
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1251)
; OTHER INFORMATION: Xaa = any amino acid, unknown, or other
US-09-778-927A-58

Query Match 95.6%; Score 109; DB 9; Length 1251;
Best Local Similarity 95.0%; Pred. No. 5e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||||||||||||||
Db        698 CCYDGACVNNDETCEQRAAR 717
```

RESULT 13
US-09-778-927A-59
; Sequence 59, Application US/09778927A
; Patent No. US20020068342A1
; GENERAL INFORMATION:
; APPLICANT: KHOSRAVI, Rami et al.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL
; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING
; FILE REFERENCE: 2786-0160P
; CURRENT APPLICATION NUMBER: US/09/778,927A
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: IL 134453
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: IL135341
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 1602
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1602)
; OTHER INFORMATION: Xaa = any amino acid, unknown, or other
US-09-778-927A-59

Query Match 95.6%; Score 109; DB 9; Length 1602;
Best Local Similarity 95.0%; Pred. No. 6.5e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||||||||||||||
Db        698 CCYDGACVNNDETCEQRAAR 717
```

RESULT 14
US-09-878-603-64
; Sequence 64, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-64

Query Match 94.7%; Score 108; DB 10; Length 20;
Best Local Similarity 95.0%; Pred. No. 8.4e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||| ||||||||||||||||
Db          1 CCYQGASVNNDETCEQRAAR 20
```

RESULT 15
US-09-878-603-65

; Sequence 65, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 65
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-65

Query Match 94.7%; Score 108; DB 10; Length 20;
Best Local Similarity 95.0%; Pred. No. 8.4e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||.|||||||||
Db          1 CCYDGASVNNQETCEQRAAR 20
```

Search completed: December 9, 2003, 14:13:18
Job time : 30 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 14:05:42 ; Search time 41 Seconds
(without alignments)
77.428 Million cell updates/sec

Title: US-09-651-685A-5
Perfect score: 114
Sequence: 1 CCYDGASVNNDETCEQRAAR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	114	100.0	20	22	AAB74055	Human C5a peptide
2	114	100.0	64	18	AAW07784	Human complement C
3	114	100.0	65	18	AAW07785	Human complement C
4	114	100.0	66	18	AAW07786	Human complement C
5	114	100.0	67	18	AAW07787	Human complement C
6	114	100.0	68	18	AAW07788	Human complement C
7	114	100.0	69	18	AAW07789	Human complement C
8	114	100.0	70	18	AAW07790	Human complement C
9	114	100.0	71	18	AAW07782	Human complement C
10	114	100.0	71	18	AAW07783	Human complement C
11	114	100.0	71	18	AAW07804	Human complement C
12	114	100.0	71	18	AAW07791	Human complement C
13	114	100.0	72	18	AAW07792	Human complement C
14	114	100.0	74	22	AAB74053	Human C5a. Homo s
15	111	97.4	20	22	AAB74111	C-terminal truncat
16	110	96.5	20	22	AAB74110	C-terminal truncat
17	110	96.5	20	22	AAB74112	C-terminal truncat
18	110	96.5	20	22	AAB74113	C-terminal truncat
19	110	96.5	20	22	AAB74114	C-terminal truncat
20	110	96.5	20	22	AAB74116	C-terminal truncat
21	110	96.5	74	22	AAB74119	Variant human C5a.
22	109	95.6	19	22	AAB74097	C-terminal truncat
23	109	95.6	20	22	AAB74120	Human C5a peptide
24	109	95.6	74	8	AAP71666	Human anaphylatoxi
25	109	95.6	74	16	AAR75497	Human C5a protein.
26	109	95.6	74	20	AAW99580	Solid phase sequen
27	109	95.6	74	22	AAE05454	Human C5a anaphyla
28	109	95.6	74	24	ABP96209	Binding epitope of
29	109	95.6	1251	23	ABG70358	Novel human thromb
30	109	95.6	1602	23	ABG70359	Novel human thromb
31	109	95.6	1676	16	AAR77604	Pro-C5 polypeptide
32	108	94.7	20	22	AAB74107	C-terminal truncat
33	108	94.7	20	22	AAB74108	C-terminal truncat
34	106	93.0	20	22	AAB74115	C-terminal truncat
35	105	92.1	18	22	AAB74098	C-terminal truncat
36	105	92.1	19	22	AAB74102	C-terminal truncat
37	102	89.5	20	22	AAB74109	C-terminal truncat
38	101	88.6	17	22	AAB74099	C-terminal truncat
39	98	86.0	20	22	AAB74117	C-terminal truncat
40	96	84.2	16	22	AAB74100	C-terminal truncat
41	96	84.2	18	22	AAB74103	C-terminal truncat
42	94	82.5	74	22	AAE05456	Cow C5a anaphylato
43	94	82.5	74	22	AAB74057	Bovine C5a. Bos s
44	93	81.6	74	22	AAE05455	Pig C5a anaphylato
45	93	81.6	74	22	AAB74058	Porcine C5a. Sus

ALIGNMENTS

RESULT 1
AAB74055
ID AAB74055 standard; Peptide; 20 AA.
XX
AC AAB74055;
XX
DT 16-MAY-2001 (first entry)
XX
DE Human C5a peptide fragment #2.
XX
KW Human; C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Homo sapiens.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
DR N-PSDB; AAF75793.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and

```
PT   treatment of sepsis in humans and other animals comprises anti-C5a
PT   antibodies generated against C-terminal truncated C5a peptides -
XX
PS   Claim 8; Page 26; 84pp; English.
XX
CC   The present sequence is a peptide fragment of human complement component
CC   C5a (the full-length sequence is given in AAB74053). The present
CC   invention relates to an antibody specific for the present sequence. The
CC   C5a-antibody can be used in a therapeutic composition, which is useful
CC   for treating a subject suffering from bacterial infection, e.g.
CC   sinusitis, meningitis, respiratory, gastrointestinal or urinary tract
CC   infections or infections in wounds. In addition, the C5a antibody can
CC   be used for treating sepsis. C5a is also known as anaphylatoxin.
XX
SQ   Sequence   20 AA;

  Query Match             100.0%;  Score 114;  DB 22;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db          1 CCYDGASVNNDETCEQRAAR 20


RESULT 2
AAW07784
ID   AAW07784 standard; protein; 64 AA.
XX
AC   AAW07784;
XX
DT   01-SEP-1997  (first entry)
XX
DE   Human complement C5a protein derivative analogue 1.
XX
KW   Human; complement; C5a; derivative; receptor; antagonist; trauma;
KW   treatment; prevention; disease; inflammation; pneumonitis; burn;
KW   adult respiratory distress syndrome; ARDS; pulmonary; injury;
KW   post myocardial; infarction; inflammatory bowel; endotoxic shock;
KW   rheumatoid arthritis; psoriasis; sepsis; transplant rejection;
KW   immunosuppressive therapy; blood transfusion; dysfunction;
KW   haemodialysis; leukopheresis; prophylaxis; reperfusion.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "wild type Thr replaced by Met"
FT   Misc-difference 27
FT                   /note= "wild type Cys replaced by Ser"
FT   Misc-difference 64
FT                   /note= "wild type Asn replaced by Cys"
XX
PN   WO9639503-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP02422.
XX
PR   05-JUN-1995;   95US-0463377.
PR   05-JUN-1995;   95US-0462648.
PR   05-JUN-1995;   95US-0463224.
XX
PA   (CIBA ) CIBA GEIGY AG.
XX
PI   Schmitz A,  Van Heeke G,  Van Oostrum J;
XX
DR   WPI; 1997-043125/04.
XX
PT   New human complement C5a poly-peptide derivs. - used as C5a receptor
PT   antagonists, partic. for treating C5a-mediated diseases and
PT   inflammatory conditions
XX
PS   Example 3; Page -; 93pp; English.
XX
CC   The present sequence, the human complement C5a derivative 1-64,
CC   Thr1Met, Cys27Ser, Asn64Cys, is a C5a receptor antagonist which
CC   exhibits no agonist activity. It can be used to treat or prevent
CC   C5a mediated diseases or inflammation, e.g. pneumonitis, adult
CC   respiratory distress syndrome (ARDS), pulmonary inflammation or
CC   injury, post myocardial infarction inflammation, inflammatory bowel
CC   disease, rheumatoid arthritis, psoriasis, endotoxic shock, sepsis,
CC   severe trauma and burns. It can also be used to treat patients
CC   suffering from transplant rejection, receiving immunosuppressive
CC   therapy or massive blood transfusion, exposed to medical devices
CC   or experiencing pulmonary dysfunction following haemodialysis or
CC   leukopheresis. It can also be used as a prophylactic, particularly
CC   in conditions caused by reperfusion, e.g. reperfusion following
CC   ischaemia, and circulatory contact with medical devices, as well as
CC   to prevent transplant rejection.
CC   Antibodies against the derivative can be used to detect or quantify
CC   the derivative and modify, e.g. neutralise, its activity in vivo.
CC   N.B. Sequence not given in specification, but constructed using the
CC   wild type sequence given on pages 51-52.
XX
SQ   Sequence   64 AA;

  Query Match             100.0%;  Score 114;  DB 18;  Length 64;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db         21 CCYDGASVNNDETCEQRAAR 40


RESULT 3
AAW07785
ID   AAW07785 standard; protein; 65 AA.
XX
AC   AAW07785;
XX
DT   01-SEP-1997  (first entry)
XX
DE   Human complement C5a protein derivative analogue 2.
XX
KW   Human; complement; C5a; derivative; receptor; antagonist; trauma;
KW   treatment; prevention; disease; inflammation; pneumonitis; burn;
KW   adult respiratory distress syndrome; ARDS; pulmonary; injury;
KW   post myocardial; infarction; inflammatory bowel; endotoxic shock;
KW   rheumatoid arthritis; psoriasis; sepsis; transplant rejection;
KW   immunosuppressive therapy; blood transfusion; dysfunction;
KW   haemodialysis; leukopheresis; prophylaxis; reperfusion.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "wild type Thr replaced by Met"
FT   Misc-difference 27
FT                   /note= "wild type Cys replaced by Ser"
FT   Misc-difference 65
FT                   /note= "wild type Ile replaced by Cys"
XX
PN   WO9639503-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP02422.
XX
PR   05-JUN-1995;   95US-0463377.
PR   05-JUN-1995;   95US-0462648.
PR   05-JUN-1995;   95US-0463224.
XX
```

PA (CIBA) CIBA GEIGY AG.
XX
PI Schmitz A, Van Heeke G, Van Oostrum J;
XX
DR WPI; 1997-043125/04.
XX
PT *New human complement* C5a poly-peptide derivs. - used as C5a receptor
PT antagonists, partic. for treating C5a-mediated diseases and
PT inflammatory conditions
XX
PS Example 3; Page -; 93pp; English.
XX
CC The present sequence, the human complement C5a derivative 1-65,
CC Thr1Met, Cys27Ser, Ile65Cys, is a C5a receptor antagonist which
CC exhibits no agonist activity. It can be used to treat or prevent
CC C5a mediated diseases or inflammation, e.g. pneumonitis, adult
CC respiratory distress syndrome (ARDS), pulmonary inflammation or
CC injury, post myocardial infarction inflammation, inflammatory bowel
CC disease, rheumatoid arthritis, psoriasis, endotoxic shock, sepsis,
CC severe trauma and burns. It can also be used to treat patients
CC suffering from transplant rejection, receiving immunosuppressive
CC therapy or massive blood transfusion, exposed to medical devices
CC or experiencing pulmonary dysfunction following haemodialysis or
CC leukopheresis. It can also be used as a prophylactic, particularly
CC in conditions caused by reperfusion, e.g. reperfusion following
CC ischaemia, and circulatory contact with medical devices, as well as
CC to prevent transplant rejection.
CC Antibodies against the derivative can be used to detect or quantify
CC the derivative and modify, e.g. neutralise, its activity in vivo.
CC N.B. Sequence not given in specification, but constructed using the
CC wild type sequence given on pages 51-52.
XX
SQ Sequence 65 AA;

Query Match 100.0%; Score 114; DB 18; Length 65;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db         21 CCYDGASVNNDETCEQRAAR 40
```

RESULT 4
AAW07786
ID AAW07786 standard; protein; 66 AA.
XX
AC AAW07786;
XX
DT 01-SEP-1997 (first entry)
XX
DE Human complement C5a protein derivative analogue 3.
XX
KW Human; complement; C5a; derivative; receptor; antagonist; trauma;
KW treatment; prevention; disease; inflammation; pneumonitis; burn;
KW adult respiratory distress syndrome; ARDS; pulmonary; injury;
KW post myocardial; infarction; inflammatory bowel; endotoxic shock;
KW rheumatoid arthritis; psoriasis; sepsis; transplant rejection;
KW immunosuppressive therapy; blood transfusion; dysfunction;
KW haemodialysis; leukopheresis; prophylaxis; reperfusion.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /note= "wild type Thr replaced by Met"
FT Misc-difference 27
FT /note= "wild type Cys replaced by Ser"
FT Misc-difference 66
FT /note= "wild type Ser replaced by Cys"
XX
PN WO9639503-A1.
XX
PD 12-DEC-1996.
XX
PF 04-JUN-1996; 96WO-EP02422.
XX
PR 05-JUN-1995; 95US-0463377.
PR 05-JUN-1995; 95US-0462648.
PR 05-JUN-1995; 95US-0463224.
XX
PA (CIBA) CIBA GEIGY AG.
XX
PI Schmitz A, Van Heeke G, Van Oostrum J;
XX
DR WPI; 1997-043125/04.
XX
PT New human complement C5a poly-peptide derivs. - used as C5a receptor
PT antagonists, partic. for treating C5a-mediated diseases and
PT inflammatory conditions
XX
PS Example 3; Page -; 93pp; English.
XX
CC The present sequence, the human complement C5a derivative 1-66,
CC Thr1Met, Cys27Ser, Ser66Cys, is a C5a receptor antagonist which
CC exhibits no agonist activity. It can be used to treat or prevent
CC C5a mediated diseases or inflammation, e.g. pneumonitis, adult
CC respiratory distress syndrome (ARDS), pulmonary inflammation or
CC injury, post myocardial infarction inflammation, inflammatory bowel
CC disease, rheumatoid arthritis, psoriasis, endotoxic shock, sepsis,
CC severe trauma and burns. It can also be used to treat patients
CC suffering from transplant rejection, receiving immunosuppressive
CC therapy or massive blood transfusion, exposed to medical devices
CC or experiencing pulmonary dysfunction following haemodialysis or
CC leukopheresis. It can also be used as a prophylactic, particularly
CC in conditions caused by reperfusion, e.g. reperfusion following
CC ischaemia, and circulatory contact with medical devices, as well as
CC to prevent transplant rejection.
CC Antibodies against the derivative can be used to detect or quantify
CC the derivative and modify, e.g. neutralise, its activity in vivo.
CC N.B. Sequence not given in specification, but constructed using the
CC wild type sequence given on pages 51-52.
XX
SQ Sequence 66 AA;

Query Match 100.0%; Score 114; DB 18; Length 66;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db         21 CCYDGASVNNDETCEQRAAR 40
```

RESULT 5
AAW07787
ID AAW07787 standard; protein; 67 AA.
XX
AC AAW07787;
XX
DT 01-SEP-1997 (first entry)
XX
DE Human complement C5a protein derivative analogue 4.
XX
KW Human; complement; C5a; derivative; receptor; antagonist; trauma;
KW treatment; prevention; disease; inflammation; pneumonitis; burn;
KW adult respiratory distress syndrome; ARDS; pulmonary; injury;
KW post myocardial; infarction; inflammatory bowel; endotoxic shock;
KW rheumatoid arthritis; psoriasis; sepsis; transplant rejection;
KW immunosuppressive therapy; blood transfusion; dysfunction;
KW haemodialysis; leukopheresis; prophylaxis; reperfusion.
XX
OS Homo sapiens.
OS Synthetic.
XX

```
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "wild type Thr replaced by Met"
FT   Misc-difference 27
FT                   /note= "wild type Cys replaced by Ser"
FT   Misc-difference 67
FT                   /note= "wild type His replaced by Cys"
XX
PN   WO9639503-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP02422.
XX
PR   05-JUN-1995;   95US-0463377.
PR   05-JUN-1995;   95US-0462648.
PR   05-JUN-1995;   95US-0463224.
XX
PA   (CIBA ) CIBA GEIGY AG.
XX
PI   Schmitz A,  Van Heeke G,  Van Oostrum J;
XX
DR   WPI; 1997-043125/04.
XX
PT   New human complement C5a poly-peptide derivs. - used as C5a receptor
PT   antagonists, partic. for treating C5a-mediated diseases and
PT   inflammatory conditions
XX
PS   Example 3; Page -; 93pp; English.
XX
CC   The present sequence, the human complement C5a derivative 1-67,
CC   Thr1Met, Cys27Ser, His67Cys, is a C5a receptor antagonist which
CC   exhibits no agonist activity. It can be used to treat or prevent
CC   C5a mediated diseases or inflammation, e.g. pneumonitis, adult
CC   respiratory distress syndrome (ARDS), pulmonary inflammation or
CC   injury, post myocardial infarction inflammation, inflammatory bowel
CC   disease, rheumatoid arthritis, psoriasis, endotoxic shock, sepsis,
CC   severe trauma and burns. It can also be used to treat patients
CC   suffering from transplant rejection, receiving immunosuppressive
CC   therapy or massive blood transfusion, exposed to medical devices
CC   or experiencing pulmonary dysfunction following haemodialysis or
CC   leukopheresis. It can also be used as a prophylactic, particularly
CC   in conditions caused by reperfusion, e.g. reperfusion following
CC   ischaemia, and circulatory contact with medical devices, as well as
CC   to prevent transplant rejection.
CC   Antibodies against the derivative can be used to detect or quantify
CC   the derivative and modify, e.g. neutralise, its activity in vivo.
CC   N.B. Sequence not given in specification, but constructed using the
CC   wild type sequence given on pages 51-52.
XX
SQ   Sequence   67 AA;

  Query Match             100.0%;  Score 114;  DB 18;  Length 67;
  Best Local Similarity   100.0%;  Pred. No. 2e-10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db         21 CCYDGASVNNDETCEQRAAR 40


RESULT 6
AAW07788
ID   AAW07788 standard; protein; 68 AA.
XX
AC   AAW07788;
XX
DT   01-SEP-1997  (first entry)
XX
DE   Human complement C5a protein derivative analogue 5.
XX
KW   Human; complement; C5a; derivative; receptor; antagonist; trauma;
KW   treatment; prevention; disease; inflammation; pneumonitis; burn;
KW   adult respiratory distress syndrome; ARDS; pulmonary; injury;
KW   post myocardial; infarction; inflammatory bowel; endotoxic shock;
KW   rheumatoid arthritis; psoriasis; sepsis; transplant rejection;
KW   immunosuppressive therapy; blood transfusion; dysfunction;
KW   haemodialysis; leukopheresis; prophylaxis; reperfusion.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "wild type Thr replaced by Met"
FT   Misc-difference 27
FT                   /note= "wild type Cys replaced by Ser"
FT   Misc-difference 68
FT                   /note= "wild type Lys replaced by Cys"
XX
PN   WO9639503-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP02422.
XX
PR   05-JUN-1995;   95US-0463377.
PR   05-JUN-1995;   95US-0462648.
PR   05-JUN-1995;   95US-0463224.
XX
PA   (CIBA ) CIBA GEIGY AG.
XX
PI   Schmitz A,  Van Heeke G,  Van Oostrum J;
XX
DR   WPI; 1997-043125/04.
XX
PT   New human complement C5a poly-peptide derivs. - used as C5a receptor
PT   antagonists, partic. for treating C5a-mediated diseases and
PT   inflammatory conditions
XX
PS   Example 3; Page -; 93pp; English.
XX
CC   The present sequence, the human complement C5a derivative 1-68,
CC   Thr1Met, Cys27Ser, Lys68Cys, is a C5a receptor antagonist which
CC   exhibits no agonist activity. It can be used to treat or prevent
CC   C5a mediated diseases or inflammation, e.g. pneumonitis, adult
CC   respiratory distress syndrome (ARDS), pulmonary inflammation or
CC   injury, post myocardial infarction inflammation, inflammatory bowel
CC   disease, rheumatoid arthritis, psoriasis, endotoxic shock, sepsis,
CC   severe trauma and burns. It can also be used to treat patients
CC   suffering from transplant rejection, receiving immunosuppressive
CC   therapy or massive blood transfusion, exposed to medical devices
CC   or experiencing pulmonary dysfunction following haemodialysis or
CC   leukopheresis. It can also be used as a prophylactic, particularly
CC   in conditions caused by reperfusion, e.g. reperfusion following
CC   ischaemia, and circulatory contact with medical devices, as well as
CC   to prevent transplant rejection.
CC   Antibodies against the derivative can be used to detect or quantify
CC   the derivative and modify, e.g. neutralise, its activity in vivo.
CC   N.B. Sequence not given in specification, but constructed using the
CC   wild type sequence given on pages 51-52.
XX
SQ   Sequence   68 AA;

  Query Match             100.0%;  Score 114;  DB 18;  Length 68;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db         21 CCYDGASVNNDETCEQRAAR 40


RESULT 7
AAW07789
ID   AAW07789 standard; protein; 69 AA.
```

XX
AC AAW07789;
XX
DT 01-SEP-1997 (first entry)
XX
DE Human complement C5a protein derivative analogue 6.
XX
KW Human; complement; C5a; derivative; receptor; antagonist; trauma;
KW treatment; prevention; disease; inflammation; pneumonitis; burn;
KW adult respiratory distress syndrome; ARDS; pulmonary; injury;
KW post myocardial; infarction; inflammatory bowel; endotoxic shock;
KW rheumatoid arthritis; psoriasis; sepsis; transplant rejection;
KW immunosuppressive therapy; blood transfusion; dysfunction;
KW haemodialysis; leukopheresis; prophylaxis; reperfusion.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /note= "wild type Thr replaced by Met"
FT Misc-difference 27
FT /note= "wild type Cys replaced by Ser"
FT Misc-difference 69
FT /note= "wild type Asp replaced by Cys"
XX
PN WO9639503-A1.
XX
PD 12-DEC-1996.
XX
PF 04-JUN-1996; 96WO-EP02422.
XX
PR 05-JUN-1995; 95US-0463377.
PR 05-JUN-1995; 95US-0462648.
PR 05-JUN-1995; 95US-0463224.
XX
PA (CIBA) CIBA GEIGY AG.
XX
PI Schmitz A, Van Heeke G, Van Oostrum J;
XX
DR WPI; 1997-043125/04.
XX
PT New human complement C5a poly-peptide derivs. - used as C5a receptor
PT antagonists, partic. for treating C5a-mediated diseases and
PT inflammatory conditions
XX
PS Example 3; Page -; 93pp; English.
XX
CC The present sequence, the human complement C5a derivative 1-69,
CC Thr1Met, Cys27Ser, Asp69Cys, is a C5a receptor antagonist which
CC exhibits no agonist activity. It can be used to treat or prevent
CC C5a mediated diseases or inflammation, e.g. pneumonitis, adult
CC respiratory distress syndrome (ARDS), pulmonary inflammation or
CC injury, post myocardial infarction inflammation, inflammatory bowel
CC disease, rheumatoid arthritis, psoriasis, endotoxic shock, sepsis,
CC severe trauma and burns. It can also be used to treat patients
CC suffering from transplant rejection, receiving immunosuppressive
CC therapy or massive blood transfusion, exposed to medical devices
CC or experiencing pulmonary dysfunction following haemodialysis or
CC leukopheresis. It can also be used as a prophylactic, particularly
CC in conditions caused by reperfusion, e.g. reperfusion following
CC ischaemia, and circulatory contact with medical devices, as well as
CC to prevent transplant rejection.
CC Antibodies against the derivative can be used to detect or quantify
CC the derivative and modify, e.g. neutralise, its activity in vivo.
CC N.B. Sequence not given in specification, but constructed using the
CC wild type sequence given on pages 51-52.
XX
SQ Sequence 69 AA;

Query Match 100.0%; Score 114; DB 18; Length 69;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||||||||||||||||
Db 21 CCYDGASVNNDETCEQRAAR 40

RESULT 8
AAW07790
ID AAW07790 standard; protein; 70 AA.
XX
AC AAW07790;
XX
DT 01-SEP-1997 (first entry)
XX
DE Human complement C5a protein derivative analogue 7.
XX
KW Human; complement; C5a; derivative; receptor; antagonist; trauma;
KW treatment; prevention; disease; inflammation; pneumonitis; burn;
KW adult respiratory distress syndrome; ARDS; pulmonary; injury;
KW post myocardial; infarction; inflammatory bowel; endotoxic shock;
KW rheumatoid arthritis; psoriasis; sepsis; transplant rejection;
KW immunosuppressive therapy; blood transfusion; dysfunction;
KW haemodialysis; leukopheresis; prophylaxis; reperfusion.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /note= "wild type Thr replaced by Met"
FT Misc-difference 27
FT /note= "wild type Cys replaced by Ser"
FT Misc-difference 70
FT /note= "wild type Met replaced by Cys"
XX
PN WO9639503-A1.
XX
PD 12-DEC-1996.
XX
PF 04-JUN-1996; 96WO-EP02422.
XX
PR 05-JUN-1995; 95US-0463377.
PR 05-JUN-1995; 95US-0462648.
PR 05-JUN-1995; 95US-0463224.
XX
PA (CIBA) CIBA GEIGY AG.
XX
PI Schmitz A, Van Heeke G, Van Oostrum J;
XX
DR WPI; 1997-043125/04.
XX
PT New human complement C5a poly-peptide derivs. - used as C5a receptor
PT antagonists, partic. for treating C5a-mediated diseases and
PT inflammatory conditions
XX
PS Example 3; Page -; 93pp; English.
XX
CC The present sequence, the human complement C5a derivative 1-70,
CC Thr1Met, Cys27Ser, Met70Cys, is a C5a receptor antagonist which
CC exhibits no agonist activity. It can be used to treat or prevent
CC C5a mediated diseases or inflammation, e.g. pneumonitis, adult
CC respiratory distress syndrome (ARDS), pulmonary inflammation or
CC injury, post myocardial infarction inflammation, inflammatory bowel
CC disease, rheumatoid arthritis, psoriasis, endotoxic shock, sepsis,
CC severe trauma and burns. It can also be used to treat patients
CC suffering from transplant rejection, receiving immunosuppressive
CC therapy or massive blood transfusion, exposed to medical devices
CC or experiencing pulmonary dysfunction following haemodialysis or
CC leukopheresis. It can also be used as a prophylactic, particularly
CC in conditions caused by reperfusion, e.g. reperfusion following
CC ischaemia, and circulatory contact with medical devices, as well as
CC to prevent transplant rejection.
CC Antibodies against the derivative can be used to detect or quantify

```
CC   the derivative and modify, e.g. neutralise, its activity in vivo.
CC   N.B. Sequence not given in specification, but constructed using the
CC   wild type sequence given on pages 51-52.
XX
SQ   Sequence   70 AA;

  Query Match             100.0%;  Score 114;  DB 18;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db         21 CCYDGASVNNDETCEQRAAR 40


RESULT 9
AAW07782
ID   AAW07782 standard; protein; 71 AA.
XX
AC   AAW07782;
XX
DT   01-SEP-1997  (first entry)
XX
DE   Human complement C5a protein derivative.
XX
KW   Human; complement; C5a; derivative; receptor; antagonist; trauma;
KW   treatment; prevention; disease; inflammation; pneumonitis; burn;
KW   adult respiratory distress syndrome; ARDS; pulmonary; injury;
KW   post myocardial; infarction; inflammatory bowel; endotoxic shock;
KW   rheumatoid arthritis; psoriasis; sepsis; transplant rejection;
KW   immunosuppressive therapy; blood transfusion; dysfunction;
KW   haemodialysis; leukopheresis; prophylaxis; reperfusion.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "wild type Thr replaced by Gly"
FT   Misc-difference 27
FT                   /note= "wild type Cys replaced by Ser"
FT   Misc-difference 71
FT                   /note= "wild type Gln replaced by Cys"
XX
PN   WO9639503-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP02422.
XX
PR   05-JUN-1995;   95US-0463377.
PR   05-JUN-1995;   95US-0462648.
PR   05-JUN-1995;   95US-0463224.
XX
PA   (CIBA ) CIBA GEIGY AG.
XX
PI   Schmitz A,  Van Heeke G,  Van Oostrum J;
XX
DR   WPI; 1997-043125/04.
XX
PT   New human complement C5a poly-peptide derivs. - used as C5a receptor
PT   antagonists, partic. for treating C5a-mediated diseases and
PT   inflammatory conditions
XX
PS   Claim 11; Page -; 93pp; English.
XX
CC   The present sequence, the human complement C5a derivative 1-71,
CC   Thr1Gly, Cys27Ser, Gln71Cys, is a C5a receptor antagonist which
CC   exhibits no agonist activity. It can be used to treat or prevent
CC   C5a mediated diseases or inflammation, e.g. pneumonitis, adult
CC   respiratory distress syndrome (ARDS), pulmonary inflammation or
CC   injury, post myocardial infarction inflammation, inflammatory bowel
CC   disease, rheumatoid arthritis, psoriasis, endotoxic shock, sepsis,
CC   severe trauma and burns. It can also be used to treat patients
CC   suffering from transplant rejection, receiving immunosuppressive
CC   therapy or massive blood transfusion, exposed to medical devices
CC   or experiencing pulmonary dysfunction following haemodialysis or
CC   leukopheresis. It can also be used as a prophylactic, particularly
CC   in conditions caused by reperfusion, e.g. reperfusion following
CC   ischaemia, and circulatory contact with medical devices, as well as
CC   to prevent transplant rejection.
CC   Antibodies against the derivative can be used to detect or quantify
CC   the derivative and modify, e.g. neutralise, its activity in vivo.
CC   N.B. Sequence not given in specification, but constructed using the
CC   wild type sequence given on pages 51-52.
XX
SQ   Sequence   71 AA;

  Query Match             100.0%;  Score 114;  DB 18;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db         21 CCYDGASVNNDETCEQRAAR 40


RESULT 10
AAW07783
ID   AAW07783 standard; protein; 71 AA.
XX
AC   AAW07783;
XX
DT   01-SEP-1997  (first entry)
XX
DE   Human complement C5a protein derivative.
XX
KW   Human; complement; C5a; derivative; receptor; antagonist; trauma;
KW   treatment; prevention; disease; inflammation; pneumonitis; burn;
KW   adult respiratory distress syndrome; ARDS; pulmonary; injury;
KW   post myocardial; infarction; inflammatory bowel; endotoxic shock;
KW   rheumatoid arthritis; psoriasis; sepsis; transplant rejection;
KW   immunosuppressive therapy; blood transfusion; dysfunction;
KW   haemodialysis; leukopheresis; prophylaxis; reperfusion.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "wild type Thr replaced by Gly"
FT   Misc-difference 27
FT                   /note= "wild type Cys replaced by Ser"
FT   Misc-difference 67
FT                   /note= "wild type His replaced by Phe"
FT   Misc-difference 71
FT                   /note= "wild type Gln replaced by Cys"
XX
PN   WO9639503-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP02422.
XX
PR   05-JUN-1995;   95US-0463377.
PR   05-JUN-1995;   95US-0462648.
PR   05-JUN-1995;   95US-0463224.
XX
PA   (CIBA ) CIBA GEIGY AG.
XX
PI   Schmitz A,  Van Heeke G,  Van Oostrum J;
XX
DR   WPI; 1997-043125/04.
XX
PT   New human complement C5a poly-peptide derivs. - used as C5a receptor
PT   antagonists, partic. for treating C5a-mediated diseases and
```

```
PT   inflammatory conditions
XX
PS   Claim 12; Page -; 93pp; English.
XX
CC   The present sequence, the human complement C5a derivative 1-71,
CC   Thr1Gly, Cys27Ser, His67Phe, Gln71Cys, is a C5a receptor antagonist
CC   which exhibits no agonist activity. It can be used to treat or
CC   prevent C5a mediated diseases or inflammation, e.g. pneumonitis,
CC   adult respiratory distress syndrome (ARDS), pulmonary inflammation
CC   or injury, post myocardial infarction inflammation, inflammatory
CC   bowel disease, rheumatoid arthritis, psoriasis, endotoxic shock,
CC   sepsis, severe trauma and burns. It can also be used to treat
CC   patients suffering from transplant rejection, receiving
CC   immunosuppressive therapy or massive blood transfusion, exposed to
CC   medical devices or experiencing pulmonary dysfunction following
CC   haemodialysis or leukopheresis. It can also be used as a
CC   prophylactic, particularly in conditions caused by reperfusion,
CC   e.g. reperfusion following ischaemia, and circulatory contact with
CC   medical devices, as well as to prevent transplant rejection.
CC   Antibodies against the derivative can be used to detect or quantify
CC   the derivative and modify, e.g. neutralise, its activity in vivo.
CC   N.B. Sequence not given in specification, but constructed using the
CC   wild type sequence given on pages 51-52.
XX
SQ   Sequence   71 AA;

  Query Match             100.0%;  Score 114;  DB 18;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db         21 CCYDGASVNNDETCEQRAAR 40


RESULT 11
AAW07804
ID   AAW07804 standard; protein; 71 AA.
XX
AC   AAW07804;
XX
DT   01-SEP-1997  (first entry)
XX
DE   Human complement C5a protein derivative.
XX
KW   Human; complement; C5a; derivative; receptor; antagonist; trauma;
KW   treatment; prevention; disease; inflammation; pneumonitis; burn;
KW   adult respiratory distress syndrome; ARDS; pulmonary; injury;
KW   post myocardial; infarction; inflammatory bowel; endotoxic shock;
KW   rheumatoid arthritis; psoriasis; sepsis; transplant rejection;
KW   immunosuppressive therapy; blood transfusion; dysfunction;
KW   haemodialysis; leukopheresis; prophylaxis; reperfusion.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "wild type Thr replaced by Met"
FT   Misc-difference 27
FT                   /note= "wild type Cys replaced by Ser"
FT   Misc-difference 67
FT                   /note= "wild type His replaced by Phe"
FT   Misc-difference 71
FT                   /note= "wild type Gln replaced by Cys"
XX
PN   WO9639503-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP02422.
XX
PR   05-JUN-1995;   95US-0463377.
PR   05-JUN-1995;   95US-0462648.
PR   05-JUN-1995;   95US-0463224.
XX
PA   (CIBA ) CIBA GEIGY AG.
XX
PI   Schmitz A,  Van Heeke G,  Van Oostrum J;
XX
DR   WPI; 1997-043125/04.
DR   N-PSDB; AAT47209.
XX
PT   New human complement C5a poly-peptide derivs. - used as C5a receptor
PT   antagonists, partic. for treating C5a-mediated diseases and
PT   inflammatory conditions
XX
PS   Example 3; Page -; 93pp; English.
XX
CC   The present sequence, the human complement C5a derivative 1-71,
CC   Thr1Met, Cys27Ser, His67Phe, Gln71Cys, is a C5a receptor antagonist
CC   which exhibits no agonist activity. It can be used to treat or
CC   prevent C5a mediated diseases or inflammation, e.g. pneumonitis,
CC   adult respiratory distress syndrome (ARDS), pulmonary inflammation
CC   or injury, post myocardial infarction inflammation, inflammatory
CC   bowel disease, rheumatoid arthritis, psoriasis, endotoxic shock,
CC   sepsis, severe trauma and burns. It can also be used to treat
CC   patients suffering from transplant rejection, receiving
CC   immunosuppressive therapy or massive blood transfusion, exposed to
CC   medical devices or experiencing pulmonary dysfunction following
CC   haemodialysis or leukopheresis. It can also be used as a
CC   prophylactic, particularly in conditions caused by reperfusion,
CC   e.g. reperfusion following ischaemia, and circulatory contact with
CC   medical devices, as well as to prevent transplant rejection.
CC   Antibodies against the derivative can be used to detect or quantify
CC   the derivative and modify, e.g. neutralise, its activity in vivo.
CC   N.B. Sequence not given in specification, but constructed using the
CC   wild type sequence given on pages 51-52.
XX
SQ   Sequence   71 AA;

  Query Match             100.0%;  Score 114;  DB 18;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db         21 CCYDGASVNNDETCEQRAAR 40


RESULT 12
AAW07791
ID   AAW07791 standard; protein; 71 AA.
XX
AC   AAW07791;
XX
DT   01-SEP-1997  (first entry)
XX
DE   Human complement C5a protein derivative analogue 8.
XX
KW   Human; complement; C5a; derivative; receptor; antagonist; trauma;
KW   treatment; prevention; disease; inflammation; pneumonitis; burn;
KW   adult respiratory distress syndrome; ARDS; pulmonary; injury;
KW   post myocardial; infarction; inflammatory bowel; endotoxic shock;
KW   rheumatoid arthritis; psoriasis; sepsis; transplant rejection;
KW   immunosuppressive therapy; blood transfusion; dysfunction;
KW   haemodialysis; leukopheresis; prophylaxis; reperfusion.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "wild type Thr replaced by Met"
FT   Misc-difference 27
FT                   /note= "wild type Cys replaced by Ser"
```

```
FT   Misc-difference 71
FT                   /note= "wild type Gln replaced by Cys"
XX
PN   WO9639503-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP02422.
XX
PR   05-JUN-1995;   95US-0463377.
PR   05-JUN-1995;   95US-0462648.
PR   05-JUN-1995;   95US-0463224.
XX
PA   (CIBA ) CIBA GEIGY AG.
XX
PI   Schmitz A,  Van Heeke G,  Van Oostrum J;
XX
DR   WPI; 1997-043125/04.
XX
PT   New human complement C5a poly-peptide derivs. - used as C5a receptor
PT   antagonists, partic. for treating C5a-mediated diseases and
PT   inflammatory conditions
XX
PS   Example 3; Page -; 93pp; English.
XX
CC   The present sequence, the human complement C5a derivative 1-71,
CC   Thr1Met, Cys27Ser, Gln71Cys, is a C5a receptor antagonist which
CC   exhibits no agonist activity. It can be used to treat or prevent
CC   C5a mediated diseases or inflammation, e.g. pneumonitis, adult
CC   respiratory distress syndrome (ARDS), pulmonary inflammation or
CC   injury, post myocardial infarction inflammation, inflammatory bowel
CC   disease, rheumatoid arthritis, psoriasis, endotoxic shock, sepsis,
CC   severe trauma and burns. It can also be used to treat patients
CC   suffering from transplant rejection, receiving immunosuppressive
CC   therapy or massive blood transfusion, exposed to medical devices
CC   or experiencing pulmonary dysfunction following haemodialysis or
CC   leukopheresis. It can also be used as a prophylactic, particularly
CC   in conditions caused by reperfusion, e.g. reperfusion following
CC   ischaemia, and circulatory contact with medical devices, as well as
CC   to prevent transplant rejection.
CC   Antibodies against the derivative can be used to detect or quantify
CC   the derivative and modify, e.g. neutralise, its activity in vivo.
CC   N.B. Sequence not given in specification, but constructed using the
CC   wild type sequence given on pages 51-52.
XX
SQ   Sequence   71 AA;

  Query Match             100.0%;  Score 114;  DB 18;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db         21 CCYDGASVNNDETCEQRAAR 40


RESULT 13
AAW07792
ID   AAW07792 standard; protein; 72 AA.
XX
AC   AAW07792;
XX
DT   01-SEP-1997  (first entry)
XX
DE   Human complement C5a protein derivative analogue 9.
XX
KW   Human; complement; C5a; derivative; receptor; antagonist; trauma;
KW   treatment; prevention; disease; inflammation; pneumonitis; burn;
KW   adult respiratory distress syndrome; ARDS; pulmonary; injury;
KW   post myocardial; infarction; inflammatory bowel; endotoxic shock;
KW   rheumatoid arthritis; psoriasis; sepsis; transplant rejection;
KW   immunosuppressive therapy; blood transfusion; dysfunction;
KW   haemodialysis; leukopheresis; prophylaxis; reperfusion.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "wild type Thr replaced by Met"
FT   Misc-difference 27
FT                   /note= "wild type Cys replaced by Ser"
FT   Misc-difference 72
FT                   /note= "wild type Leu replaced by Cys"
XX
PN   WO9639503-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-EP02422.
XX
PR   05-JUN-1995;   95US-0463377.
PR   05-JUN-1995;   95US-0462648.
PR   05-JUN-1995;   95US-0463224.
XX
PA   (CIBA ) CIBA GEIGY AG.
XX
PI   Schmitz A,  Van Heeke G,  Van Oostrum J;
XX
DR   WPI; 1997-043125/04.
XX
PT   New human complement C5a poly-peptide derivs. - used as C5a receptor
PT   antagonists, partic. for treating C5a-mediated diseases and
PT   inflammatory conditions
XX
PS   Example 3; Page -; 93pp; English.
XX
CC   The present sequence, the human complement C5a derivative 1-72,
CC   Thr1Met, Cys27Ser, Leu72Cys, is a C5a receptor antagonist which
CC   exhibits no agonist activity. It can be used to treat or prevent
CC   C5a mediated diseases or inflammation, e.g. pneumonitis, adult
CC   respiratory distress syndrome (ARDS), pulmonary inflammation or
CC   injury, post myocardial infarction inflammation, inflammatory bowel
CC   disease, rheumatoid arthritis, psoriasis, endotoxic shock, sepsis,
CC   severe trauma and burns. It can also be used to treat patients
CC   suffering from transplant rejection, receiving immunosuppressive
CC   therapy or massive blood transfusion, exposed to medical devices
CC   or experiencing pulmonary dysfunction following haemodialysis or
CC   leukopheresis. It can also be used as a prophylactic, particularly
CC   in conditions caused by reperfusion, e.g. reperfusion following
CC   ischaemia, and circulatory contact with medical devices, as well as
CC   to prevent transplant rejection.
CC   Antibodies against the derivative can be used to detect or quantify
CC   the derivative and modify, e.g. neutralise, its activity in vivo.
CC   N.B. Sequence not given in specification, but constructed using the
CC   wild type sequence given on pages 51-52.
XX
SQ   Sequence   72 AA;

  Query Match             100.0%;  Score 114;  DB 18;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||||||||||||||||
Db         21 CCYDGASVNNDETCEQRAAR 40


RESULT 14
AAB74053
ID   AAB74053 standard; protein; 74 AA.
XX
AC   AAB74053;
XX
DT   16-MAY-2001  (first entry)
XX
```

DE Human C5a.
XX
KW Human; C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS *Homo sapiens*.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
DR N-PSDB; AAF75791.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Example 7; Page 26; 84pp; English.
XX
CC The present sequence is human complement component C5a. The present
CC invention relates to an antibody specific for the present sequence. The
CC C5a-antibody can be used in a therapeutic composition, which is useful
CC for treating a subject suffering from bacterial infection, e.g.
CC sinusitis, meningitis, respiratory, gastrointestinal or urinary tract
CC infections or infections in wounds. In addition, the C5a antibody can
CC be used for treating sepsis. C5a is also known as anaphylatoxin.
XX
SQ Sequence 74 AA;

Query Match 100.0%; Score 114; DB 22; Length 74;
Best Local Similarity 100.0%; Pred. No. 2.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||||||||||||||||
Db 21 CCYDGASVNNDETCEQRAAR 40

RESULT 15
AAB74111
ID AAB74111 standard; Peptide; 20 AA.
XX
AC AAB74111;
XX
DT 16-MAY-2001 (first entry)
XX
DE C-terminal truncated C5a peptide #49.
XX
KW C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Unidentified.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Disclosure; Page 30; 84pp; English.
XX
CC The present sequence is a C-terminal truncated C5a peptide fragment. The
CC present invention relates to an antibody specific for the present
CC sequence. The C5a-antibody can be used in a therapeutic composition,
CC which is useful for treating a subject suffering from bacterial
CC infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC urinary tract infections or infections in wounds. In addition, the C5a
CC antibody can be used for treating sepsis. C5a is also known as
CC anaphylatoxin.
XX
SQ Sequence 20 AA;

Query Match 97.4%; Score 111; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.6e-10;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||:|||||||||||||
Db 1 CCYDGATVNNDETCEQRAAR 20

Search completed: December 9, 2003, 14:06:39
Job time : 42 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 14:05:43 ; Search time 35 Seconds
(without alignments)
147.459 Million cell updates/sec

Title: US-09-651-685A-5
Perfect score: 114
Sequence: 1 CCYDGASVNNDETCEQRAAR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	84	73.7	77	11	Q63078	Q63078 rattus norv
2	84	73.7	742	11	Q8BNV3	Q8bnv3 mus musculu
3	81	71.1	348	11	Q8BWN9	Q8bwn9 mus musculu
4	81	71.1	352	11	Q8BWS8	Q8bws8 mus musculu
5	56.5	49.6	1148	11	Q61372	Q61372 mus musculu
6	56.5	49.6	1738	11	O70346	O70346 mus musculu
7	51	44.7	608	5	Q9VI74	Q9vi74 drosophila
8	50.5	44.3	1614	13	Q98977	Q98977 oncorhynchu
9	49.5	43.4	483	5	Q9N497	Q9n497 caenorhabdi
10	48	42.1	135	12	Q65686	Q65686 beet necrot
11	48	42.1	282	12	Q9DGS4	Q9dgs4 beet necrot
12	48	42.1	282	12	Q9DGS5	Q9dgs5 beet necrot
13	48	42.1	282	12	Q65670	Q65670 beet necrot
14	48	42.1	282	12	Q9DS67	Q9ds67 beet necrot
15	48	42.1	282	12	Q9DS85	Q9ds85 beet necrot
16	48	42.1	282	12	Q9DS75	Q9ds75 beet necrot
17	47.5	41.7	1684	13	Q9DDV9	Q9ddv9 oncorhynchu
18	47	41.2	227	5	Q9GTY7	Q9gty7 plasmodium
19	47	41.2	834	12	Q85429	Q85429 rice stripe
20	47	41.2	834	12	Q85430	Q85430 rice stripe
21	47	41.2	1625	12	O55597	O55597 garlic viru
22	46	40.4	265	16	Q8XH54	Q8xh54 clostridium
23	45.5	39.9	1159	13	Q9YIA6	Q9yia6 cyprinus ca
24	45.5	39.9	1640	13	Q9YIA8	Q9yia8 cyprinus ca
25	45.5	39.9	1642	13	Q9YIA9	Q9yia9 cyprinus ca
26	45.5	39.9	1642	13	Q9YIB0	Q9yib0 cyprinus ca
27	45	39.5	141	16	Q8FJ17	Q8fj17 escherichia
28	45	39.5	159	16	Q9CGL1	Q9cgl1 lactococcus
29	45	39.5	163	12	Q9J813	Q9j813 spodoptera
30	45	39.5	221	5	Q9NL25	Q9nl25 plasmodium
31	45	39.5	226	5	Q9N6J1	Q9n6j1 plasmodium
32	45	39.5	226	5	Q9NL21	Q9nl21 plasmodium
33	45	39.5	226	5	Q9N6G5	Q9n6g5 plasmodium
34	45	39.5	231	5	Q9N673	Q9n673 plasmodium
35	45	39.5	231	5	Q9NL20	Q9nl20 plasmodium
36	45	39.5	231	5	Q9N681	Q9n681 plasmodium
37	45	39.5	231	5	Q9NL24	Q9nl24 plasmodium
38	45	39.5	231	5	Q9NL22	Q9nl22 plasmodium
39	45	39.5	231	5	Q9NL19	Q9nl19 plasmodium
40	45	39.5	231	5	Q9NL23	Q9nl23 plasmodium
41	45	39.5	236	5	Q9NL18	Q9nl18 plasmodium
42	45	39.5	236	5	O96556	O96556 plasmodium
43	45	39.5	282	12	Q9DS77	Q9ds77 beet necrot
44	45	39.5	394	5	Q93521	Q93521 caenorhabdi
45	45	39.5	739	2	O87381	O87381 haemophilus

ALIGNMENTS

RESULT 1
Q63078
ID Q63078 PRELIMINARY; PRT; 77 AA.
AC Q63078;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE C5a complement component protein (Fragment).
GN C5A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Lewis; TISSUE=Liver;
RX MEDLINE=97236424; PubMed=9116048;
RA Rothermel E., Rolf O., Goetze O., Zwirner J.;
RT "Nucleotide and corrected amino acid sequence of the functional
RT recombinant rat anaphylatoxin C5a.";
RL Biochim. Biophys. Acta 1351:9-12(1997).
DR EMBL; X91892; CAA62994.1; -.
DR HSSP; P01031; 1KJS.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR001840; Anaphylatoxn.
DR Pfam; PF01821; ANATO; 1.
DR PRINTS; PR00004; ANAPHYLATOXN.
DR ProDom; PD003264; Anaphylatoxin; 1.
DR SMART; SM00104; ANATO; 1.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
FT NON_TER 1 1
FT NON_TER 77 77
SQ SEQUENCE 77 AA; 8981 MW; 14141F41CC38BD28 CRC64;

Query Match 73.7%; Score 84; DB 11; Length 77;
Best Local Similarity 75.0%; Pred. No. 1.2e-06;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||  |  ||||||  ||
Db         24 CCYDGARENKYETCEQRVAR 43


RESULT 2
Q8BNV3
ID   Q8BNV3      PRELIMINARY;      PRT;   742 AA.
AC   Q8BNV3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolytic complement (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Aorta and vein;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK080166; BAC37836.1; -.
FT   NON_TER     742    742
SQ   SEQUENCE   742 AA;  83511 MW;  C730312767BA043D CRC64;

  Query Match             73.7%;  Score 84;  DB 11;  Length 742;
  Best Local Similarity   75.0%;  Pred. No. 1.4e-05;
  Matches   15;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||  ||  ||||:|  ||
Db        702 CCYDGARVNFYETCEERVAR 721


RESULT 3
Q8BWN9
ID   Q8BWN9      PRELIMINARY;      PRT;   348 AA.
AC   Q8BWN9;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical Anaphylotoxins.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK050413; BAC34242.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   348 AA;  39591 MW;  A27D35A4D5B5E19D CRC64;

  Query Match             71.1%;  Score 81;  DB 11;  Length 348;
  Best Local Similarity   75.0%;  Pred. No. 2e-05;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||  |  |||  |||||
Db        276 CCYDGARYNIHETCVQRAAR 295


RESULT 4
Q8BWS8
ID   Q8BWS8      PRELIMINARY;      PRT;   352 AA.
AC   Q8BWS8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical Anaphylotoxins.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK050126; BAC34079.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   352 AA;  40039 MW;  174A853E21CEFE11 CRC64;

  Query Match             71.1%;  Score 81;  DB 11;  Length 352;
  Best Local Similarity   75.0%;  Pred. No. 2e-05;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||  |  |||  |||||
Db        295 CCYDGARYNIHETCVQRAAR 314


RESULT 5
Q61372
ID   Q61372      PRELIMINARY;      PRT;  1148 AA.
AC   Q61372;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   C4 complement protein (Fragment).
GN   C4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B10.D2(C4(h)Slp(a));
RX   MEDLINE=86176748; PubMed=3008092;
RA   Hemenway C., Kalff M., Stavenhagen J., Walthall D., Robins D.;
RT   "Sequence comparison of alleles of the fourth component of complement
RT   (C4) and sex-limited protein (Slp).";
RL   Nucleic Acids Res. 14:2539-2554(1986).
DR   EMBL; X05314; CAA28936.1; -.
DR   HSSP; P01031; 1KJS.
DR   MGD; MGI:88228; C4.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
```

```
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   1148 AA;  127361 MW;  2710ECF832B6F0C9 CRC64;

  Query Match             49.6%;  Score 56.5;  DB 11;  Length 1148;
  Best Local Similarity   57.1%;  Pred. No. 1;
  Matches   12;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          1 CCYDG-ASVNNDETCEQRAAR 20
              || ||    :    ||||||||
Db        110 CCQDGMTKLPMKRTCEQRAAR 130


RESULT 6
O70346
ID   O70346      PRELIMINARY;      PRT;  1738 AA.
AC   O70346;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Complement C4.
GN   STK19 OR C4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rowen L., Qin S., Lasky S.R., Loretz C., Dors M., Mahairas G.,
RA   Hood L.;
RT   "Sequence of the mouse major histocompatibility locus class III
RT   region.";
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF049850; AAC05279.1; -.
DR   HSSP; P01031; 1KJS.
DR   MGD; MGI:88228; C4.
DR   MGD; MGI:1860085; Stk19.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
SQ   SEQUENCE   1738 AA;  192944 MW;  BDDB2802091EEE81 CRC64;

  Query Match             49.6%;  Score 56.5;  DB 11;  Length 1738;
  Best Local Similarity   57.1%;  Pred. No. 1.6;
  Matches   12;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          1 CCYDG-ASVNNDETCEQRAAR 20
              || ||    :    ||||||||
Db        700 CCQDGMTKLPMKRTCEQRAAR 720


RESULT 7
Q9VI74
ID   Q9VI74      PRELIMINARY;      PRT;   608 AA.
AC   Q9VI74;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG10055 protein.
GN   CG10055.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003672; AAF54051.1; -.
DR   FlyBase; FBgn0037482; CG10055.
DR   InterPro; IPR007009; SHQ1.
DR   Pfam; PF04925; SHQ1; 1.
SQ   SEQUENCE   608 AA;  69145 MW;  035736E5F2567014 CRC64;

  Query Match             44.7%;  Score 51;  DB 5;  Length 608;
  Best Local Similarity   57.1%;  Pred. No. 4.5;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNNDETCE 15
              |||  : ||: |||
Db        327 CYDVRTTNNEPTCE 340


RESULT 8
Q98977
ID   Q98977      PRELIMINARY;      PRT;  1614 AA.
AC   Q98977;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Complement component C3-3 (Fragment).
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
```

```
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96323260; PubMed=8710907;
RA   Sunyer J.O., Zarkadis I.K., Sahu A., Lambris J.D.;
RT   "Multiple forms of complement C3 in trout that differ in binding to
RT   complement activators.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:8546-8551(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=20437839; PubMed=10980316;
RA   Zarkadis I.K., Sarrias M.R., Sfyroera G., Sunyer J.O., Lambris J.D.;
RT   "Cloning and structure of three rainbow trout C3 molecules: a
RT   plausible explanation for their functional diversity.";
RL   Dev. Comp. Immunol. 25:11-24(2001).
DR   EMBL; U61753; AAC60015.2; -.
DR   HSSP; P01024; 1C3D.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   1614 AA;  180714 MW;  0910CE3B73D38F10 CRC64;

  Query Match             44.3%;  Score 50.5;  DB 13;  Length 1614;
  Best Local Similarity   57.9%;  Pred. No. 16;
  Matches   11;  Conservative    1;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 CCYDGASVN-NDETCEQRA 18
              || ||   |  | |||:||
Db        640 CCMDGMRKNILDYTCERRA 658


RESULT 9
Q9N497
ID   Q9N497      PRELIMINARY;      PRT;   483 AA.
AC   Q9N497;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein Y110A2AL.2.
GN   Y110A2AL.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Waterston R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Smith A.;
RT   "The sequence of C. elegans cosmid Y110A2AL.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC024746; AAF60400.2; -.
KW   Hypothetical protein.
SQ   SEQUENCE   483 AA;  52777 MW;  6633184625D9B360 CRC64;

  Query Match             43.4%;  Score 49.5;  DB 5;  Length 483;
  Best Local Similarity   47.4%;  Pred. No. 6.3;
  Matches    9;  Conservative    3;  Mismatches    2;  Indels    5;  Gaps    1;

Qy          1 CC-----YDGASVNNDETC 14
              ||     :||:| | :|||
Db        133 CCAKRDVFDGSSSNRNETC 151


RESULT 10
Q65686
ID   Q65686      PRELIMINARY;      PRT;   135 AA.
AC   Q65686;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 15.0 kDa protein.
OS   Beet necrotic yellow vein virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
OX   NCBI_TaxID=31721;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=F2;
RA   Bouzoubaa S., Guilley H., Jonard G., Richards K., Putz C.;
RT   "nucleotide sequence analysis of rna-3 and rna-4 of beet necrotic
RT   yellow vein virus, isolates f2 and g1.";
RL   J. Gen. Virol. 66:1553-1564(1985).
DR   EMBL; M36896; AAA69656.1; -.
DR   InterPro; IPR007004; DUF656.
DR   Pfam; PF04920; DUF656; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   135 AA;  15027 MW;  79E2231BA5240226 CRC64;

  Query Match             42.1%;  Score 48;  DB 12;  Length 135;
  Best Local Similarity   57.1%;  Pred. No. 2.9;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNNDETCE 15
              || |  | :|| ||
Db        115 CYSGVDVLSDELCE 128


RESULT 11
Q9DGS4
ID   Q9DGS4      PRELIMINARY;      PRT;   282 AA.
AC   Q9DGS4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 32.1 kDa protein.
OS   Beet necrotic yellow vein virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
OX   NCBI_TaxID=31721;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=F75, and F72;
RX   MEDLINE=20456802; PubMed=11003469;
RA   Koenig R., Lennefors B.L.;
RT   "Molecular analyses of European A, B and P type sources of Beet
RT   necrotic yellow vein virus and detection of the rare P type in
RT   Kazakhstan.";
RL   Arch. Virol. 145:1561-1570(2000).
```

```
DR   EMBL; AF197548; AAG37084.1; -.
DR   EMBL; AF197546; AAG37078.1; -.
DR   InterPro; IPR007004; DUF656.
DR   Pfam; PF04920; DUF656; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   282 AA;  32053 MW;  F8DB465FBC6B5C1B CRC64;

  Query Match             42.1%;  Score 48;  DB 12;  Length 282;
  Best Local Similarity   57.1%;  Pred. No. 6.4;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNNDETCE 15
              || |   | :|| ||
Db        115 CYSGVDVLSDELCE 128


RESULT 12
Q9DGS5
ID   Q9DGS5      PRELIMINARY;      PRT;   282 AA.
AC   Q9DGS5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 32.1 kDa protein.
OS   Beet necrotic yellow vein virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
OX   NCBI_TaxID=31721;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Kas3, and Kas2;
RX   MEDLINE=20456802; PubMed=11003469;
RA   Koenig R., Lennefors B.L.;
RT   "Molecular analyses of European A, B and P type sources of Beet
RT   necrotic yellow vein virus and detection of the rare P type in
RT   Kazakhstan.";
RL   Arch. Virol. 145:1561-1570(2000).
DR   EMBL; AF197557; AAG37097.1; -.
DR   EMBL; AF197554; AAG37090.1; -.
DR   InterPro; IPR007004; DUF656.
DR   Pfam; PF04920; DUF656; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   282 AA;  32111 MW;  F8D90D3208FB5C1B CRC64;

  Query Match             42.1%;  Score 48;  DB 12;  Length 282;
  Best Local Similarity   57.1%;  Pred. No. 6.4;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNNDETCE 15
              || |   | :|| ||
Db        115 CYSGVDVLSDELCE 128


RESULT 13
Q65670
ID   Q65670      PRELIMINARY;      PRT;   282 AA.
AC   Q65670;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   31K protein.
OS   Beet necrotic yellow vein virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
OX   NCBI_TaxID=31721;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S;
RX   MEDLINE=97128991; PubMed=8973531;
RA   Saito M., Kiguchi T., Kusume T., Tamada T.;
RT   "Complete nucleotide sequence of the Japanese isolate S of beet
RT   necrotic yellow vein virus RNA and comparison with European
RT   isolates.";
RL   Arch. Virol. 141:2163-2175(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C;
RA   Yu J., Han C., Yang L., Li D., Liu Y.;
RT   "cDNA cloning, sequencing and expression of RNA4 from beet necrotic
RT   yellow vein virus.";
RL   Acta Microbiol. Sin. 37:7-14(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C;
RA   Li D.;
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; D84413; BAA12347.1; -.
DR   EMBL; AJ239199; CAA15427.1; -.
DR   InterPro; IPR007004; DUF656.
DR   Pfam; PF04920; DUF656; 1.
SQ   SEQUENCE   282 AA;  31970 MW;  D96AA3B079E063E6 CRC64;

  Query Match             42.1%;  Score 48;  DB 12;  Length 282;
  Best Local Similarity   57.1%;  Pred. No. 6.4;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNNDETCE 15
              || |   | :|| ||
Db        115 CYSGVDVLSDELCE 128


RESULT 14
Q9DS67
ID   Q9DS67      PRELIMINARY;      PRT;   282 AA.
AC   Q9DS67;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 31.9 kDa protein.
OS   Beet necrotic yellow vein virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
OX   NCBI_TaxID=31721;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N7;
RX   MEDLINE=20456802; PubMed=11003469;
RA   Koenig R., Lennefors B.L.;
RT   "Molecular analyses of European A, B and P type sources of Beet
RT   necrotic yellow vein virus and detection of the rare P type in
RT   Kazakhstan.";
RL   Arch. Virol. 145:1561-1570(2000).
DR   EMBL; AF197559; AAG37099.1; -.
DR   InterPro; IPR007004; DUF656.
DR   Pfam; PF04920; DUF656; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   282 AA;  31948 MW;  58A34878C868131F CRC64;

  Query Match             42.1%;  Score 48;  DB 12;  Length 282;
  Best Local Similarity   57.1%;  Pred. No. 6.4;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNNDETCE 15
              || |   | :|| ||
Db        115 CYSGVDVLSDELCE 128


RESULT 15
Q9DS85
ID   Q9DS85      PRELIMINARY;      PRT;   282 AA.
AC   Q9DS85;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 31.9 kDa protein.
OS   Beet necrotic yellow vein virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
OX   NCBI_TaxID=31721;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=D71;
RX   MEDLINE=20456802; PubMed=11003469;
RA   Koenig R., Lennefors B.L.;
RT   "Molecular analyses of European A, B and P type sources of Beet
RT   necrotic yellow vein virus and detection of the rare P type in
RT   Kazakhstan.";
RL   Arch. Virol. 145:1561-1570(2000).
DR   EMBL; AF197544; AAG37076.1; -.
DR   InterPro; IPR007004; DUF656.
DR   Pfam; PF04920; DUF656; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   282 AA;  31939 MW;  3383394A3A739F2D CRC64;

  Query Match             42.1%;  Score 48;  DB 12;  Length 282;
  Best Local Similarity   57.1%;  Pred. No. 6.4;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNNDETCE 15
              || |   | :|| ||
Db        115 CYSGVDVLSDELCE 128


Search completed: December  9, 2003, 14:07:28
Job time : 37 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 14:05:43 ; Search time 11 Seconds
(without alignments)
85.503 Million cell updates/sec

Title: US-09-651-685A-5
Perfect score: 114
Sequence: 1 CCYDGASVNNDETCEQRAAR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	109	95.6	1676	1	CO5_HUMAN	P01031 homo sapien
2	94	82.5	74	1	CO5A_BOVIN	P12082 bos taurus
3	93	81.6	74	1	CO5A_PIG	P01032 sus scrofa
4	84	73.7	76	1	CO5A_RAT	P08650 rattus norv
5	84	73.7	1680	1	CO5_MOUSE	P06684 mus musculu
6	56.5	49.6	1738	1	CO4_MOUSE	P01029 mus musculu
7	55.5	48.7	920	1	CO4_BOVIN	P01030 bos taurus
8	55.5	48.7	1737	1	CO4_RAT	P08649 rattus norv
9	51.5	45.2	1744	1	CO4_HUMAN	P01028 homo sapien
10	48	42.1	282	1	Y32K_BNYVG	P19231 beet necrot
11	46	40.4	1473	1	OVOS_CHICK	P20740 gallus gall
12	44.5	39.0	281	1	Y125_MYCPN	P75511 mycoplasma
13	44.5	39.0	685	1	CFAH_BOVIN	Q28085 bos taurus
14	44.5	39.0	1249	1	UN36_CAEEL	P34374 caenorhabdi
15	44	38.6	624	1	ARLY_VIBPA	P59620 vibrio para
16	44	38.6	624	1	ARLY_VIBVU	Q8dcm9 vibrio vuln
17	44	38.6	715	1	S141_HUMAN	Q92503 homo sapien
18	44	38.6	2095	1	RRPL_TOSV	P37800 toscana vir
19	43.5	38.2	1651	1	CO3_NAJNA	Q01833 naja naja (
20	43	37.7	337	1	DNJ2_CAEEL	Q17433 caenorhabdi
21	43	37.7	461	1	SYC_YERPE	Q8zcc0 yersinia pe
22	42.5	37.3	705	1	FBL1_MOUSE	Q08879 mus musculu
23	42.5	37.3	1640	1	CO3_ONCMY	P98093 oncorhynchu
24	42.5	37.3	1673	1	CO3_LAMJA	Q00685 lampetra ja
25	42	36.8	241	1	TSN1_HUMAN	O60635 homo sapien
26	42	36.8	261	1	RFA4_HUMAN	Q13156 homo sapien
27	42	36.8	263	1	FSL3_HUMAN	O95633 homo sapien
28	42	36.8	760	1	AD25_MOUSE	Q9r159 mus musculu
29	42	36.8	1068	1	YCF0_MARPO	P12221 marchantia
30	42	36.8	2149	1	RRPL_RVFVZ	P27316 rift valley
31	41.5	36.4	65	1	MT_PARLI	P80367 paracentrot
32	41.5	36.4	486	1	MTN3_HUMAN	O15232 homo sapien
33	41	36.0	332	1	KC2A_MAIZE	P28523 zea mays (m
34	41	36.0	450	1	VD10_BPT5	P11107 bacteriopha
35	41	36.0	491	1	YO84_MYCTU	O53209 mycobacteri
36	41	36.0	735	1	AD02_MACFA	Q28478 macaca fasc
37	41	36.0	1826	1	SUIS_RABIT	P07768 oryctolagus
38	41	36.0	2230	1	GOG4_HUMAN	Q13439 homo sapien
39	41	36.0	4543	1	LRP1_CHICK	P98157 gallus gall
40	41	36.0	4660	1	LRP2_RAT	P98158 rattus norv
41	40	35.1	351	1	Y4VJ_RHISN	Q53218 rhizobium s
42	40	35.1	366	1	YGDE_ECOLI	P32066 escherichia
43	40	35.1	448	1	EXG1_YEAST	P23776 saccharomyc
44	40	35.1	448	1	GIAN_DROME	P39572 drosophila
45	40	35.1	503	1	AMPA_RHILO	Q984s1 rhizobium l

ALIGNMENTS

RESULT 1
CO5_HUMAN
ID CO5_HUMAN STANDARD; PRT; 1676 AA.
AC P01031;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Complement C5 precursor [Contains: C5a anaphylatoxin].
GN C5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91079575; PubMed=1984448;
RA Haviland D.L., Haviland J.C., Fleischer D.T., Hunt A., Wetsel R.A.;
RT "Complete cDNA sequence of human complement pro-C5. Evidence of
RT truncated transcripts derived from a single copy gene.";
RL J. Immunol. 146:362-368(1991).
RN [2]
RP SEQUENCE OF 412-1676 FROM N.A.
RX MEDLINE=88209511; PubMed=3365401;
RA Wetsel R.A., Lemons R.S., Lebeau M.M., Barnum S.R., Noack D.,
RA Tack B.F.;
RT "Molecular analysis of human complement component C5: localization of
RT the structural gene to chromosome 9.";
RL Biochemistry 27:1474-1482(1988).
RN [3]
RP SEQUENCE OF 412-902 FROM N.A.
RX MEDLINE=85130937; PubMed=2579066;
RA Lundwall A.B., Wetsel R.A., Kristensen T., Whitehead A.S.,
RA Woods D.E., Ogden R.C., Colten H.R., Tack B.F.;
RT "Isolation and sequence analysis of a cDNA clone encoding the fifth
RT complement component.";
RL J. Biol. Chem. 260:2108-2112(1985).
RN [4]
RP SEQUENCE OF 678-751.
RX MEDLINE=79005687; PubMed=690134;
RA Fernandez H.N., Hugli T.E.;
RT "Primary structural analysis of the polypeptide portion of human C5a
RT anaphylatoxin. Polypeptide sequence determination and assignment of
RT the oligosaccharide attachment site in C5a.";
RL J. Biol. Chem. 253:6955-6964(1978).
RN [5]
RP SEQUENCE OF 678-751 FROM N.A.
RX MEDLINE=91144547; PubMed=1996961;
RA Bohnsack J.F., Mollison K.W., Buko A.M., Ashworth J.C., Hill H.R.;
RT "Group B streptococci inactivate complement component C5a by enzymic
RT cleavage at the C-terminus.";
RL Biochem. J. 273:635-640(1991).
RN [6]
RP STRUCTURE BY NMR OF C5A.
RX MEDLINE=88309754; PubMed=3408713;
RA Zuiderweg E.R.P., Mollison K.W., Henkin J., Carter G.W.;
RT "Sequence-specific assignments in the 1H NMR spectrum of the human

```
RT   inflammatory protein C5a.";
RL   Biochemistry 27:3568-3580(1988).
RN   [7]
RP   STRUCTURE BY NMR OF C5A.
RX   MEDLINE=89207527; PubMed=2784981;
RA   Zuiderweg E.R.P., Nettesheim D.G., Mollison K.W., Carter G.W.;
RT   "Tertiary structure of human complement component C5a in solution
RT   from nuclear magnetic resonance data.";
RL   Biochemistry 28:172-185(1989).
RN   [8]
RP   STRUCTURE BY NMR OF C5A.
RX   MEDLINE=89274164; PubMed=2730871;
RA   Zuiderweg E.R.P., Fesik S.W.;
RT   "Heteronuclear three-dimensional NMR spectroscopy of the inflammatory
RT   protein C5a.";
RL   Biochemistry 28:2387-2391(1989).
RN   [9]
RP   STRUCTURE BY NMR OF C5A.
RX   MEDLINE=97160477; PubMed=9007977;
RA   Zhang X., Boyar W., Galakatos N., Gonnella N.C.;
RT   "Solution structure of a unique C5a semi-synthetic antagonist:
RT   implications in receptor binding.";
RL   Protein Sci. 6:65-72(1997).
RN   [10]
RP   STRUCTURE BY NMR OF C5A.
RX   MEDLINE=97332508; PubMed=9188742;
RA   Zhang X., Boyar W., Toth M.J., Wennogle L., Gonnella N.C.;
RT   "Structural definition of the C5a C terminus by two-dimensional
RT   nuclear magnetic resonance spectroscopy.";
RL   Proteins 28:261-267(1997).
CC   -!- FUNCTION: ACTIVATION OF C5 BY A C5 CONVERTASE INITIATES THE
CC       SPONTANEOUS ASSEMBLY OF THE LATE COMPLEMENT COMPONENTS, C5-C9,
CC       INTO THE MEMBRANE ATTACK COMPLEX. C5B HAS A TRANSIENT BINDING SITE
CC       FOR C6. THE C5B-C6 COMPLEX IS THE FOUNDATION UPON WHICH THE LYTIC
CC       COMPLEX IS ASSEMBLED.
CC   -!- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C5,
CC       C5 ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC       INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC       PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC       BASOPHILIC LEUKOCYTES. C5A ALSO STIMULATES THE LOCOMOTION OF
CC       POLYMORPHONUCLEAR LEUKOCYTES (CHEMOKINESIS) AND DIRECT THEIR
CC       MIGRATION TOWARD SITES OF INFLAMMATION (CHEMOTAXIS).
CC   -!- SUBUNIT: C5 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 BASIC
CC       RESIDUES, FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE
CC       BOND. C5 CONVERTASE ACTIVATES C5 BY CLEAVING THE ALPHA CHAIN,
CC       RELEASING C5A ANAPHYLATOXIN & GENERATING C5B (BETA CHAIN + ALPHA'
CC       CHAIN).
CC   -!- SIMILARITY: TO C3, C4 AND ALPHA-2-MACROGLOBULIN.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 855
CC       ONWARD DUE TO THE PRESENCE OF AN ALU REPEAT.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M57729; AAA51925.1; -.
DR   EMBL; M65134; AAA51856.1; -.
DR   PIR; A40075; C5HU.
DR   PDB; 1KJS; 15-MAY-97.
DR   PDB; 1CFA; 17-SEP-97.
DR   Genew; HGNC:1331; C5.
DR   MIM; 120900; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0003797; F:antibacterial peptide activity; TAS.
DR   GO; GO:0008009; F:chemokine activity; TAS.
DR   GO; GO:0005102; F:receptor binding activity; TAS.
DR   GO; GO:0000187; P:activation of MAPK; TAS.
DR   GO; GO:0006935; P:chemotaxis; TAS.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
DR   GO; GO:0006954; P:inflammatory response; TAS.
DR   GO; GO:0009618; P:response to pathogenic bacteria; TAS.
DR   GO; GO:0006950; P:response to stress; TAS.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; MacrogloblnA2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; FALSE_NEG.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Complement alternate pathway; Glycoprotein;
KW   Plasma; Membrane attack complex; Cytolysis; Inflammatory response;
KW   Signal; Polymorphism; 3D-structure.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    673       COMPLEMENT C5 BETA CHAIN.
FT   PROPEP      674    677
FT   CHAIN       678   1676       COMPLEMENT C5 ALPHA CHAIN.
FT   PEPTIDE     678    751       C5A ANAPHYLATOXIN.
FT   CHAIN       752   1676       C5B (ALPHA').
FT   DOMAIN      698    732       ANAPHYLATOXIN-LIKE.
FT   DISULFID    698    724
FT   DISULFID    699    731
FT   DISULFID    711    732
FT   CARBOHYD    741    741       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    911    911       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1115   1115       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1630   1630       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     518    518       F -> S.
FT                                /FTId=VAR_001996.
FT   VARIANT     802    802       I -> V (IN dbSNP:17611).
FT                                /FTId=VAR_014574.
FT   VARIANT    1053   1053       M -> L (IN dbSNP:17609).
FT                                /FTId=VAR_014575.
FT   VARIANT    1310   1310       S -> N (IN dbSNP:17610).
FT                                /FTId=VAR_014576.
FT   VARIANT    1437   1437       E -> D (IN dbSNP:17612).
FT                                /FTId=VAR_014577.
FT   TURN        679    680
FT   HELIX       682    688
FT   TURN        689    689
FT   HELIX       693    703
FT   TURN        704    704
FT   HELIX       711    715
FT   TURN        716    716
FT   HELIX       722    739
FT   HELIX       745    748
FT   TURN        749    750
SQ   SEQUENCE   1676 AA;  188331 MW;  87DCAA65FF977D19 CRC64;

  Query Match             95.6%;  Score 109;  DB 1;  Length 1676;
  Best Local Similarity   95.0%;  Pred. No. 2.8e-09;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||||||||||||||
Db        698 CCYDGACVNNDETCEQRAAR 717


RESULT 2
CO5A_BOVIN
ID   CO5A_BOVIN     STANDARD;      PRT;    74 AA.
AC   P12082;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
```

```
DE   Complement C5A anaphylatoxin.
GN   C5.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=86136134; PubMed=3081348;
RA   Gennaro R., Simonic T., Negri A., Mottola C., Secchi C., Ronchi S.,
RA   Romeo D.;
RT   "C5a fragment of bovine complement. Purification, bioassays,
RT   amino-acid sequence and other structural studies.";
RL   Eur. J. Biochem. 155:77-86(1986).
RN   [2]
RP   STRUCTURE BY NMR.
RX   MEDLINE=89005703; PubMed=3262536;
RA   Zarbock J., Gennaro R., Romeo D., Clore G.M., Gronenborn A.M.;
RT   "A proton nuclear magnetic resonance study of the conformation of
RT   bovine anaphylatoxin C5a in solution.";
RL   FEBS Lett. 238:289-294(1988).
CC   -!- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C5,
CC       C5 ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC       INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC       PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC       BASOPHILIC LEUKOCYTES. C5A ALSO STIMULATES THE LOCOMOTION OF
CC       POLYMORPHONUCLEAR LEUKOCYTES (CHEMOKINESIS) AND DIRECT THEIR
CC       MIGRATION TOWARD SITES OF INFLAMMATION (CHEMOTAXIS).
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
DR   PIR; A25408; A25408.
DR   HSSP; P01032; 1C5A.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR001599; MacrogloblnA2.
DR   Pfam; PF01821; ANATO; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; PARTIAL.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Complement alternate pathway; Plasma;
KW   Inflammatory response.
FT   DOMAIN       21     55       ANAPHYLATOXIN-LIKE.
FT   DISULFID     21     47       BY SIMILARITY.
FT   DISULFID     22     54       BY SIMILARITY.
FT   DISULFID     34     55       BY SIMILARITY.
SQ   SEQUENCE   74 AA;  8517 MW;  C09DF742D12D70F6 CRC64;

  Query Match             82.5%;  Score 94;  DB 1;  Length 74;
  Best Local Similarity   80.0%;  Pred. No. 2.4e-08;
  Matches   16;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||  |:|||||:||||
Db         21 CCYDGAHRNDDETCEERAAR 40


RESULT 3
CO5A_PIG
ID   CO5A_PIG       STANDARD;      PRT;    74 AA.
AC   P01032;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Complement C5A anaphylatoxin.
GN   C5.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=80182137; PubMed=7372604;
RA   Gerard C., Hugli T.E.;
RT   "Amino acid sequence of the anaphylatoxin from the fifth component of
RT   porcine complement.";
RL   J. Biol. Chem. 255:4710-4715(1980).
RN   [2]
RP   ACTIVE REGION.
RX   MEDLINE=81199549; PubMed=6940191;
RA   Gerard C., Hugli T.E.;
RT   "Identification of classical anaphylatoxin as the des-Arg form of the
RT   C5a molecule: evidence of a modulator role for the oligosaccharide
RT   unit in human des-Arg74-C5a.";
RL   Proc. Natl. Acad. Sci. U.S.A. 78:1833-1837(1981).
RN   [3]
RP   STRUCTURE BY NMR.
RX   MEDLINE=90248365; PubMed=2337573;
RA   Williamson M.P., Madison V.S.;
RT   "Three-dimensional structure of porcine C5adesArg from 1H nuclear
RT   magnetic resonance data.";
RL   Biochemistry 29:2895-2905(1990).
CC   -!- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C5,
CC       C5 ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC       INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC       PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC       BASOPHILIC LEUKOCYTES. C5A ALSO STIMULATES THE LOCOMOTION OF
CC       POLYMORPHONUCLEAR LEUKOCYTES (CHEMOKINESIS) AND DIRECT THEIR
CC       MIGRATION TOWARD SITES OF INFLAMMATION (CHEMOTAXIS).
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
DR   PIR; A01268; A01268.
DR   PDB; 1C5A; 15-OCT-91.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; MacrogloblnA2.
DR   Pfam; PF01821; ANATO; 1.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; PARTIAL.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Complement alternate pathway; Plasma;
KW   Inflammatory response; 3D-structure.
FT   DOMAIN       21     55       ANAPHYLATOXIN-LIKE.
FT   DISULFID     21     47
FT   DISULFID     22     54
FT   DISULFID     34     55
FT   SITE         72     74       ARE REQUIRED FOR 90% OF C5A ACTIVITY
FT                                (ALTHOUGH ARG-74 IS NOT ESSENTIAL).
FT   HELIX         2     11
FT   TURN         13     14
FT   HELIX        16     26
FT   HELIX        34     40
FT   HELIX        45     62
FT   TURN         63     64
SQ   SEQUENCE   74 AA;  8609 MW;  11AAF2E94A026EB3 CRC64;

  Query Match             81.6%;  Score 93;  DB 1;  Length 74;
  Best Local Similarity   80.0%;  Pred. No. 3.5e-08;
  Matches   16;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||  |:|||||:||||
Db         21 CCYDGAYRNDDETCEERAAR 40


RESULT 4
CO5A_RAT
ID   CO5A_RAT       STANDARD;      PRT;    76 AA.
AC   P08650;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Complement C5A anaphylatoxin.
GN   C5.
```

```
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE.
RA   Cui L.-X., Ferreri K., Hugli T.E.;
RT   "Characterization of rat C5a, a uniquely active spasmogen.";
RL   Complement 2:18-19(1985).
CC   -!- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C5,
CC       C5 ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC       INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC       PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC       BASOPHILIC LEUKOCYTES. C5A ALSO STIMULATES THE LOCOMOTION OF
CC       POLYMORPHONUCLEAR LEUKOCYTES (CHEMOKINESIS) AND DIRECT THEIR
CC       MIGRATION TOWARD SITES OF INFLAMMATION (CHEMOTAXIS).
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
DR   HSSP; P01031; 1KJS.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; MacroglobInA2.
DR   Pfam; PF01821; ANATO; 1.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; PARTIAL.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Complement alternate pathway; Glycoprotein;
KW   Plasma; Inflammatory response.
FT   DOMAIN       24     58       ANAPHYLATOXIN-LIKE.
FT   DISULFID     24     50       BY SIMILARITY.
FT   DISULFID     25     57       BY SIMILARITY.
FT   DISULFID     37     58       BY SIMILARITY.
FT   CARBOHYD     66     66       N-LINKED (GLCNAC. . .).
SQ   SEQUENCE   76 AA;  8869 MW;  2EC15183A6E54769 CRC64;

  Query Match             73.7%;  Score 84;  DB 1;  Length 76;
  Best Local Similarity   75.0%;  Pred. No. 9.8e-07;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||  |  |||||| ||
Db         24 CCYDGARENKYETCEQRVAR 43


RESULT 5
CO5_MOUSE
ID   CO5_MOUSE      STANDARD;      PRT;  1680 AA.
AC   P06684;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Complement C5 precursor (Hemolytic complement) [Contains: C5A
DE   anaphylatoxin].
GN   C5 OR HC.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90153853; PubMed=2303408;
RA   Wetsel R.A., Fleischer D.T., Haviland D.L.;
RT   "Deficiency of the murine fifth complement component (C5). A 2-base
RT   pair gene deletion in a 5'-exon.";
RL   J. Biol. Chem. 265:2435-2440(1990).
RN   [2]
RP   SEQUENCE OF 41-1680 FROM N.A.
RX   MEDLINE=87185363; PubMed=2436653;
RA   Wetsel R.A., Ogata R.T., Tack B.F.;
RT   "Primary structure of the fifth component of murine complement.";
RL   Biochemistry 26:737-743(1987).
CC   -!- FUNCTION: ACTIVATION OF C5 BY A C5 CONVERTASE INITIATES THE
CC       SPONTANEOUS ASSEMBLY OF THE LATE COMPLEMENT COMPONENTS, C5-C9,
CC       INTO THE MEMBRANE ATTACK COMPLEX. C5B HAS A TRANSIENT BINDING SITE
CC       FOR C6. THE C5B-C6 COMPLEX IS THE FOUNDATION UPON WHICH THE LYTIC
CC       COMPLEX IS ASSEMBLED.
CC   -!- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C5,
CC       C5 ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC       INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC       PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC       BASOPHILIC LEUKOCYTES. C5A ALSO STIMULATES THE LOCOMOTION OF
CC       POLYMORPHONUCLEAR LEUKOCYTES (CHEMOKINESIS) AND DIRECT THEIR
CC       MIGRATION TOWARD SITES OF INFLAMMATION (CHEMOTAXIS).
CC   -!- SUBUNIT: C5 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 BASIC
CC       RESIDUES, FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE
CC       BOND. C5 CONVERTASE ACTIVATES C5 BY CLEAVING THE ALPHA CHAIN,
CC       RELEASING C5A ANAPHYLATOXIN & GENERATING C5B (BETA CHAIN + ALPHA'
CC       CHAIN).
CC   -!- SIMILARITY: TO C3, C4 AND ALPHA-2-MACROGLOBULIN.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M35525; AAA37349.1; -.
DR   EMBL; M35526; AAA37348.1; -.
DR   PIR; A35530; C5MS.
DR   HSSP; P01031; 1KJS.
DR   MGD; MGI:96031; Hc.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; FALSE_NEG.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Complement alternate pathway; Glycoprotein;
KW   Plasma; Membrane attack complex; Cytolysis; Inflammatory response;
KW   Signal.
FT   SIGNAL       1     18
FT   CHAIN       19   1680       COMPLEMENT C5.
FT   CHAIN       19    674       COMPLEMENT C5 BETA CHAIN.
FT   PROPEP     675    678
FT   CHAIN      679   1680       COMPLEMENT C5 ALPHA CHAIN.
FT   PEPTIDE    679    755       C5A ANAPHYLATOXIN.
FT   CHAIN      756   1680       C5B (ALPHA').
FT   DOMAIN     702    736       ANAPHYLATOXIN-LIKE.
FT   DISULFID   702    728       BY SIMILARITY.
FT   DISULFID   703    735       BY SIMILARITY.
FT   DISULFID   715    736       BY SIMILARITY.
FT   CARBOHYD   427    427       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   915    915       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1119   1119       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1633   1633       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT    216    216       Y -> L (IN DEFECTIVE VARIANT C5D).
FT   VARIANT    217   1680       MISSING (IN DEFECTIVE VARIANT C5D).
SQ   SEQUENCE  1680 AA;  188877 MW;  81EB5A16FAC7D95C CRC64;

  Query Match             73.7%;  Score 84;  DB 1;  Length 1680;
  Best Local Similarity   75.0%;  Pred. No. 2.7e-05;
  Matches   15;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||  ||||:| ||
Db        702 CCYDGARVNFYETCEERVAR 721


RESULT 6
CO4_MOUSE
ID   CO4_MOUSE      STANDARD;      PRT;  1738 AA.
AC   P01029; Q61859;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Complement C4 precursor [Contains: C4A anaphylatoxin].
GN   C4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85298264; PubMed=3862104;
RA   Sepich D.S., Noonan D.J., Ogata R.T.;
RT   "Complete cDNA sequence of the fourth component of murine
RT   complement.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:5895-5899(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B12.WR;
RX   MEDLINE=87309760; PubMed=3624865;
RA   Rosa P.A., Sepich D.S., Robins D.M., Ogata R.T.;
RT   "Constitutive expression of Slp genes in mouse strain B10.WR directed
RT   by C4 regulatory sequences.";
RL   J. Immunol. 139:1568-1577(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B12.WR; TISSUE=Liver;
RX   MEDLINE=89380278; PubMed=2777798;
RA   Ogata R.T., Rosa P.A., Zepf N.E.;
RT   "Sequence of the gene for murine complement component C4.";
RL   J. Biol. Chem. 264:16565-16572(1989).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FM; TISSUE=Liver;
RX   MEDLINE=85289294; PubMed=2993295;
RA   Nonaka M., Nakayama K., Yeul Y.D., Takahashi M.;
RT   "Complete nucleotide and derived amino acid sequences of the fourth
RT   component of mouse complement (C4). Evolutionary aspects.";
RL   J. Biol. Chem. 260:10936-10943(1985).
RN   [5]
RP   SEQUENCE OF 651-810 AND 924-1083 FROM N.A.
RX   MEDLINE=85038607; PubMed=6208559;
RA   Nonaka M., Takahashi M., Natsuume-Sakai S., Nonaka M., Tanaka S.,
RA   Shimizu A., Honjo T.;
RT   "Isolation of cDNA clones specifying the fourth component of mouse
RT   complement and its isotype, sex-limited protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:6822-6826(1984).
RN   [6]
RP   SEQUENCE OF 1099-1142 FROM N.A.
RC   STRAIN=B10.WR, C57BL/6, C3H/He, CBA/J, B10.BR, and DBA2;
RX   MEDLINE=90353398; PubMed=2387317;
RA   Ogata R.T., Zepf N.E.;
RT   "C4 from C4-high and C4-low mouse strains have identical sequences in
RT   the region corresponding to the isotype-specific segment of human
RT   C4.";
RL   Eur. J. Immunol. 20:1607-1610(1990).
RN   [7]
RP   SEQUENCE OF 1105-1449 FROM N.A.
RX   MEDLINE=85166208; PubMed=3856857;
RA   Levi-Strauss M., Tosi M., Steinmetz M., Klein J., Meo T.;
RT   "Multiple duplications of complement C4 gene correlate with H-2-
RT   controlled testosterone-independent expression of its sex-limited
RT   isoform, C4-Slp.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:1746-1750(1985).
RN   [8]
RP   SEQUENCE OF 1257-1376 FROM N.A.
RX   MEDLINE=85038859; PubMed=6149581;
RA   Tosi M., Levi-Strauss M., Duponchel C., Meo T.;
RT   "Sequence heterogeneity of murine complementary DNA clones related to
RT   the C4 and C4-Slp isoforms of the fourth complement component.";
RL   Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:389-394(1984).
RN   [9]
RP   SEQUENCE OF 1360-1511 FROM N.A.
RX   MEDLINE=83273751; PubMed=6192448;
RA   Ogata R.T., Shreffler D.C., Sepich D.S., Lilly S.P.;
RT   "cDNA clone spanning the alpha-gamma subunit junction in the
RT   precursor of the murine fourth complement component (C4).";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:5061-5065(1983).
RN   [10]
RP   SEQUENCE OF 1-128 FROM N.A.
RC   STRAIN=FM; TISSUE=Liver;
RX   MEDLINE=86031969; PubMed=2997024;
RA   Nonaka M., Nakayama K., Yeul Y.D., Shimizu A., Takahashi M.;
RT   "Molecular cloning and characterization of complementary and genomic
RT   DNA clones for mouse C4 and Slp.";
RL   Immunol. Rev. 87:81-99(1985).
RN   [11]
RP   SEQUENCE OF 1-21 FROM N.A.
RX   MEDLINE=87017050; PubMed=3464002;
RA   Nonaka M., Kimura H., Yeul Y.D., Yokoyama S., Nakayama K.,
RA   Takahashi M.;
RT   "Identification of the 5'-flanking regulatory region responsible for
RT   the difference in transcriptional control between mouse complement C4
RT   and Slp genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:7883-7887(1986).
CC   -!- FUNCTION: C4 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC       CLASSICAL PATHWAY OF THE COMPLEMENT SYSTEM. IT IS PROCESSED BY
CC       ACTIVATED C1 WHICH REMOVE FROM THE ALPHA CHAIN THE C4A
CC       ANAPHYLATOXIN.
CC   -!- SUBUNIT: THIS PROTEIN IS SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR
CC       AND, PRIOR TO SECRETION, IS ENZYMATICALLY CLEAVED TO FORM A TRIMER
CC       OF NONIDENTICAL CHAINS (ALPHA, BETA, AND GAMMA).
CC   -!- MISCELLANEOUS: C4 IS A MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-III
CC       PROTEIN.
CC   -!- SIMILARITY: TO C3, C5 AND ALPHA-2-MACROGLOBULIN.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K00019; AAA39554.1; -.
DR   EMBL; M11729; AAA39506.1; -.
DR   EMBL; M12968; AAA39558.1; -.
DR   EMBL; M12970; AAA39555.1; -.
DR   EMBL; M12972; AAA39556.1; -.
DR   EMBL; M12969; AAA39559.1; -.
DR   EMBL; M11789; AAA39557.1; -.
DR   EMBL; K02798; AAC42021.1; -.
DR   EMBL; M17440; AAA39561.1; -.
DR   EMBL; M14225; AAA39563.1; -.
DR   EMBL; M14226; AAA39684.1; -.
DR   EMBL; X55493; CAA39112.1; -.
DR   EMBL; X55495; CAA39114.1; -.
DR   PIR; A24558; A24558.
DR   PIR; A29176; A29176.
DR   HSSP; P01031; 1KJS.
DR   MGD; MGI:88228; C4.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; Macroglobln_A2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
```

```
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Plasma; Glycoprotein; Sulfation; Signal;
KW   Inflammatory response; Thioester bond.
FT   SIGNAL       1    19
FT   CHAIN       20   673       COMPLEMENT C4, BETA CHAIN.
FT   PROPEP     674   677
FT   CHAIN      678  1443       COMPLEMENT C4, ALPHA CHAIN.
FT   PROPEP    1444  1447
FT   CHAIN     1448  1738       COMPLEMENT C4, GAMMA CHAIN.
FT   PEPTIDE    678   753       C4A ANAPHYLATOXIN.
FT   DOMAIN     700   734       ANAPHYLATOXIN-LIKE.
FT   DISULFID   700   726       BY SIMILARITY.
FT   DISULFID   701   733       BY SIMILARITY.
FT   DISULFID   714   734       BY SIMILARITY.
FT   CROSSLNK  1006  1009       Isoglutamyl cysteine thioester (Cys-Gln)
FT                              (By similarity).
FT   MOD_RES   1413  1413       SULFATION.
FT   MOD_RES   1416  1416       SULFATION.
FT   MOD_RES   1417  1417       SULFATION.
FT   CARBOHYD   224   224       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   743   743       N-LINKED (GLCNAC. . .).
FT   CARBOHYD  1387  1387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT   132   132       F -> Y (IN REF. 4).
FT   CONFLICT   327   327       G -> E (IN REF. 4).
FT   CONFLICT   570   570       Q -> E (IN REF. 4).
FT   CONFLICT   720   720       R -> G (IN REF. 5).
FT   CONFLICT   739   740       DL -> AI (IN REF. 5).
FT   CONFLICT   838   838       P -> R (IN REF. 4).
FT   CONFLICT   993   993       P -> L (IN REF. 5).
FT   CONFLICT  1043  1043       D -> E (IN REF. 5).
FT   CONFLICT  1119  1119       V -> A (IN REF. 7).
FT   CONFLICT  1190  1190       A -> T (IN REF. 7).
FT   CONFLICT  1324  1324       K -> N (IN REF. 4).
FT   CONFLICT  1401  1401       G -> S (IN REF. 9).
FT   CONFLICT  1442  1442       R -> K (IN REF. 4).
FT   CONFLICT  1453  1453       A -> V (IN REF. 4).
SQ   SEQUENCE  1738 AA;  192870 MW;  D1EE02AE7AB42BFF CRC64;

  Query Match             49.6%;  Score 56.5;  DB 1;  Length 1738;
  Best Local Similarity   57.1%;  Pred. No. 0.69;
  Matches   12;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          1 CCYDG-ASVNNDETCEQRAAR 20
              || ||   :   ||||||||
Db        700 CCQDGMTKLPMKRTCEQRAAR 720


RESULT 7
CO4_BOVIN
ID   CO4_BOVIN      STANDARD;      PRT;   920 AA.
AC   P01030; Q27992; Q27993;
DT   21-JUL-1986 (Rel. 01, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Complement C4 precursor [Contains: C4A anaphylatoxin] (Fragments).
GN   C4.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE OF 1-77.
RX   MEDLINE=83126436; PubMed=6760852;
RA   Smith M.A., Gerrie L.M., Dunbar B., Fothergill J.E.;
RT   "Primary structure of bovine complement activation fragment C4a, the
RT   third anaphylatoxin. Purification and complete amino acid sequence.";
RL   Biochem. J. 207:253-260(1982).
RN   [2]
RP   SEQUENCE OF 78-920 FROM N.A.
RC   TISSUE=Liver;
RA   Groth D.M.;
RL   Submitted (OCT-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: C4 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC       CLASSICAL PATHWAY OF THE COMPLEMENT SYSTEM. IT IS PROCESSED BY
CC       ACTIVATED C1 WHICH REMOVE FROM THE ALPHA CHAIN THE C4A
CC       ANAPHYLATOXIN.
CC   -!- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C4,
CC       C4A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC       INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC       PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC       BASOPHILIC LEUKOCYTES.
CC   -!- SUBUNIT: THIS PROTEIN IS SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR
CC       AND, PRIOR TO SECRETION, IS ENZYMATICALLY CLEAVED TO FORM A TRIMER
CC       OF NONIDENTICAL CHAINS (ALPHA, BETA, AND GAMMA).
CC   -!- MISCELLANEOUS: C4 IS A MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-III
CC       PROTEIN.
CC   -!- SIMILARITY: TO C3, C5 AND ALPHA-2-MACROGLOBULIN.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U16750; AAA52751.1; -.
DR   EMBL; U16749; AAA52750.1; -.
DR   PIR; A01265; A01265.
DR   HSSP; P01031; 1CFA.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; Macroglobln_A2.
DR   InterPro; IPR001134; Netrin_C.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Plasma; Glycoprotein; Sulfation;
KW   Inflammatory response; Thioester bond.
FT   NON_TER      1     1
FT   NON_CONS    77    78
FT   NON_CONS   342   343
FT   CHAIN        1   622       COMPLEMENT C4, ALPHA CHAIN.
FT   PROPEP     623   629
FT   CHAIN      630   920       COMPLEMENT C4, GAMMA CHAIN.
FT   PEPTIDE      1    77       C4A ANAPHYLATOXIN.
FT   DOMAIN      23    57       ANAPHYLATOXIN-LIKE.
FT   DISULFID    23    49       BY SIMILARITY.
FT   DISULFID    24    56       BY SIMILARITY.
FT   DISULFID    37    57       BY SIMILARITY.
FT   CROSSLNK   191   194       Isoglutamyl cysteine thioester (Cys-Gln).
FT   MOD_RES    593   593       SULFATION (BY SIMILARITY).
FT   MOD_RES    596   596       SULFATION (BY SIMILARITY).
FT   MOD_RES    598   598       SULFATION (BY SIMILARITY).
FT   CARBOHYD   504   504       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   567   567       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   920 AA;  101881 MW;  1425A0BD945F5497 CRC64;

  Query Match             48.7%;  Score 55.5;  DB 1;  Length 920;
  Best Local Similarity   57.1%;  Pred. No. 0.5;
  Matches   12;  Conservative    2;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 CCYDGAS-VNNDETCEQRAAR 20
              || || : :   ||||||||
```

```
Db         23 CCQDGLTRLPMARTCEQRAAR 43


RESULT 8
CO4_RAT
ID   CO4_RAT        STANDARD;      PRT;  1737 AA.
AC   P08649; Q62895; Q8R403;
DT   01-AUG-1988 (Rel. 08, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Complement C4 precursor [Contains: C4A anaphylatoxin].
GN   C4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RA   Chen C.-B., Wallis R.;
RT   "Substrate recognition by zymogen and activated forms of mannose-
RT   binding protein-associated serine proteases.";
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-266 FROM N.A.
RC   STRAIN=Brown Norway;
RX   MEDLINE=22131572; PubMed=12136338;
RA   Walter L., Hurt P., Himmelbauer H., Sudbrak R., Guenther E.;
RT   "Physical mapping of the major histocompatibility complex class II and
RT   class III regions of the rat.";
RL   Immunogenetics 54:268-275(2002).
RN   [3]
RP   SEQUENCE OF 1656-1737 FROM N.A., AND INDUCTION.
RC   STRAIN=Sprague-Dawley; TISSUE=Hepatic stellate cells;
RX   MEDLINE=96399092; PubMed=8805663;
RA   Fimmel C.J., Brown K.E., O'Neill R., Kladney R.D.;
RT   "Complement C4 protein expression by rat hepatic stellate cells.";
RL   J. Immunol. 157:2601-2609(1996).
RN   [4]
RP   SEQUENCE OF 678-753.
RX   MEDLINE=88334568; PubMed=3262196;
RA   Cui L.-X., Ferreri K., Hugli T.E.;
RT   "Structural characterization of the C4a anaphylatoxin from rat.";
RL   Mol. Immunol. 25:663-671(1988).
RN   [5]
RP   SEQUENCE OF 678-753.
RA   Cui L.-X., Ferreri K., Hugli T.E.;
RT   "Characterization of rat anaphylatoxins C4a and C5a.";
RL   Fed. Proc. 44:991-991(1985).
CC   -!- FUNCTION: C4 plays a central role in the activation of the
CC       classical pathway of the complement system. It is processed by
CC       activated C1 which removes from the alpha chain the C4A
CC       anaphylatoxin.
CC   -!- FUNCTION: Derived from proteolytic degradation of complement C4,
CC       C4A anaphylatoxin is a mediator of local inflammatory processes.
CC       It induces the contraction of smooth muscle, increases vascular
CC       permeability and causes histamine release from mast cells and
CC       basophilic leukocytes.
CC   -!- SUBUNIT: This protein is synthesized as a single-chain precursor
CC       and, prior to secretion, is enzymatically cleaved to form a trimer
CC       of nonidentical chains (alpha, beta and gamma).
CC   -!- INDUCTION: Induced in hepatic stellate cells by iron overload and
CC       by gamma-interferon.
CC   -!- MISCELLANEOUS: C4 is a major histocompatibility complex class-III
CC       protein.
CC   -!- SIMILARITY: TO C3, C5 AND ALPHA-2-MACROGLOBULIN.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   -!- SIMILARITY: Contains 1 NTR domain.
CC   -!- CAUTION: Ref.3 sequence differs from that shown due to a
CC       frameshift in position 1721.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY149995; AAN72415.1; -.
DR   EMBL; AY091787; AAM14719.1; -.
DR   EMBL; U42719; AAA91231.1; ALT_FRAME.
DR   PIR; JL0036; JL0036.
DR   HSSP; P01031; 1KJS.
DR   GO; GO:0005576; C:extracellular; IDA.
DR   GO; GO:0003811; F:complement activity activity; NAS.
DR   GO; GO:0006958; P:complement activation, classical pathway; NAS.
DR   GO; GO:0006954; P:inflammatory response; NAS.
DR   GO; GO:0045987; P:positive regulation of smooth muscle contra. . .; NAS.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF01821; ANATO; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Glycoprotein; Sulfation; Signal;
KW   Inflammatory response; Thioester bond.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    673       COMPLEMENT C4, BETA CHAIN (BY
FT                                 SIMILARITY).
FT   PROPEP      674    677       BY SIMILARITY.
FT   CHAIN       678   1442       COMPLEMENT C4, ALPHA CHAIN (BY
FT                                 SIMILARITY).
FT   PROPEP     1443   1446       BY SIMILARITY.
FT   CHAIN      1447   1737       COMPLEMENT C4, GAMMA CHAIN (BY
FT                                 SIMILARITY).
FT   PEPTIDE     678    753       C4A ANAPHYLATOXIN.
FT   DOMAIN      700    734       ANAPHYLATOXIN-LIKE.
FT   DOMAIN     1583   1735       NTR.
FT   DISULFID    700    726       BY SIMILARITY.
FT   DISULFID    701    733       BY SIMILARITY.
FT   DISULFID    714    734       BY SIMILARITY.
FT   CROSSLNK   1005   1008       Isoglutamyl cysteine thioester (Cys-Gln)
FT                                 (By similarity).
FT   MOD_RES    1412   1412       SULFATION (BY SIMILARITY).
FT   MOD_RES    1414   1414       SULFATION (POTENTIAL).
FT   MOD_RES    1416   1416       SULFATION (BY SIMILARITY).
FT   MOD_RES    1676   1676       SULFATION (POTENTIAL).
FT   CARBOHYD    224    224       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    664    664       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    743    743       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   1323   1323       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1386   1386       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    706    706       T -> A (IN REF. 5).
FT   CONFLICT   1700   1700       T -> S (IN REF. 3).
FT   CONFLICT   1709   1709       R -> H (IN REF. 3).
FT   CONFLICT   1731   1731       S -> R (IN REF. 3).
SQ   SEQUENCE   1737 AA;  192161 MW;  67FA7BFA27A3DDFA CRC64;

  Query Match             48.7%;  Score 55.5;  DB 1;  Length 1737;
  Best Local Similarity   57.1%;  Pred. No. 1;
  Matches   12;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          1 CCYDG-ASVNNDETCEQRAAR 20
              || ||   :    ||||||||
Db        700 CCQDGMTKLPMARTCEQRAAR 720


RESULT 9
```

```
CO4_HUMAN
ID   CO4_HUMAN      STANDARD;      PRT;  1744 AA.
AC   P01028; Q13160; Q13906; Q14835; Q9NPK5; Q9UIP5;
DT   21-JUL-1986 (Rel. 01, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Complement C4 precursor [Contains: C4A anaphylatoxin].
GN   C4A AND C4B.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (C4A AND C4B).
RC   TISSUE=Liver;
RX   MEDLINE=84156544; PubMed=6546707;
RA   Belt K.T., Carroll M.C., Porter R.R.;
RT   "The structural basis of the multiple forms of human complement
RT   component C4.";
RL   Cell 36:907-914(1984).
RN   [2]
RP   SEQUENCE FROM N.A. (C4A).
RX   MEDLINE=91108039; PubMed=1988494;
RA   Yu C.Y.;
RT   "The complete exon-intron structure of a human complement component
RT   C4A gene. DNA sequences, polymorphism, and linkage to the
RT   21-hydroxylase gene.";
RL   J. Immunol. 146:1057-1066(1991).
RN   [3]
RP   SEQUENCE FROM N.A. (C4B).
RC   TISSUE=Blood;
RX   MEDLINE=96163032; PubMed=8575831;
RA   Ulgiati D., Townend D.C., Christiansen F.T., Dawkins R.L.,
RA   Abraham L.J.;
RT   "Complete sequence of the complement C4 gene from the HLA-A1, B8,
RT   C4AQ0, C4B1, DR3 haplotype.";
RL   Immunogenetics 43:250-252(1996).
RN   [4]
RP   SEQUENCE FROM N.A. (C4B).
RA   Rowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,
RA   Banta A., Swartzell S., Smith T.M., Spies T., Hood L.;
RT   "Sequence determination of 300 kilobases of the human class III MHC
RT   locus.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A. (C4A).
RA   Barlow K.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE OF 1-22 AND 1056-1225 FROM N.A.
RX   MEDLINE=85156269; PubMed=3838531;
RA   Belt K.T., Yu C.Y., Carroll M.C., Porter R.R.;
RT   "Polymorphism of human complement component C4.";
RL   Immunogenetics 21:173-180(1985).
RN   [7]
RP   SEQUENCE OF 680-756.
RX   MEDLINE=81264286; PubMed=6167582;
RA   Moon K.E., Gorski J.P., Hugli T.E.;
RT   "Complete primary structure of human C4a anaphylatoxin.";
RL   J. Biol. Chem. 256:8685-8692(1981).
RN   [8]
RP   SEQUENCE OF 957-1044.
RX   MEDLINE=82182029; PubMed=6978711;
RA   Campbell R.D., Gagnon J., Porter R.R.;
RT   "Amino acid sequence around the thiol and reactive acyl groups of
RT   human complement component C4.";
RL   Biochem. J. 199:359-370(1981).
RN   [9]
RP   SEQUENCE OF 990-1037.
RX   MEDLINE=82150875; PubMed=6950384;
RA   Harrison R.A., Thomas M.L., Tack B.F.;
RT   "Sequence determination of the thiolester site of the fourth
RT   component of human complement.";
RL   Proc. Natl. Acad. Sci. U.S.A. 78:7388-7392(1981).
RN   [10]
RP   SEQUENCE OF 1-21 FROM N.A.
RX   MEDLINE=94282044; PubMed=8012361;
RA   Sargent C.A., Anderson M.J., Hsieh S.L., Kendall E.,
RA   Gomez-Escobar N., Campbell R.D.;
RT   "Characterisation of the novel gene G11 lying adjacent to the
RT   complement C4A gene in the human major histocompatibility complex.";
RL   Hum. Mol. Genet. 3:481-488(1994).
RN   [11]
RP   SEQUENCE OF 1405-1431, AND SULFATION.
RX   MEDLINE=86111851; PubMed=3944109;
RA   Hortin G., Sims H., Strauss A.W.;
RT   "Identification of the site of sulfation of the fourth component of
RT   human complement.";
RL   J. Biol. Chem. 261:1786-1793(1986).
RN   [12]
RP   STRUCTURAL BASIS OF POLYMORPHISM.
RX   MEDLINE=87080272; PubMed=2431902;
RA   Yu C.Y., Belt K.T., Giles C.M., Campbell R.D., Porter R.R.;
RT   "Structural basis of the polymorphism of human complement components
RT   C4A and C4B: gene size, reactivity and antigenicity.";
RL   EMBO J. 5:2873-2881(1986).
RN   [13]
RP   VARIANT C4A6 ALLOTYPE.
RX   MEDLINE=92242905; PubMed=1573268;
RA   Anderson M.J., Milner C.M., Cotton G.H., Campbell R.D.;
RT   "The coding sequence of the hemolytically inactive C4A6 allotype of
RT   human complement component C4 reveals that a single arginine to
RT   tryptophan substitution at beta-chain residue 458 is the likely cause
RT   of the defect.";
RL   J. Immunol. 148:2795-2802(1992).
CC   -!- FUNCTION: C4 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC       CLASSICAL PATHWAY OF THE COMPLEMENT SYSTEM. IT IS PROCESSED BY
CC       ACTIVATED C1 WHICH REMOVE FROM THE ALPHA CHAIN THE C4A
CC       ANAPHYLATOXIN.
CC   -!- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C4,
CC       C4A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC       INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC       PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC       BASOPHILIC LEUKOCYTES.
CC   -!- SUBUNIT: THIS PROTEIN IS SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR
CC       AND, PRIOR TO SECRETION, IS ENZYMATICALLY CLEAVED TO FORM A TRIMER
CC       OF NONIDENTICAL CHAINS (ALPHA, BETA, AND GAMMA).
CC   -!- POLYMORPHISM: HUMAN COMPLEMENT COMPONENT C4 IS POLYMORPHIC WITH AT
CC       LEAST TWO LOCI, C4A AND C4B. 13 ALLELES OF C4A AND 22 ALLELES OF
CC       C4B HAVE BEEN DETECTED. THE ALLELE SHOWN HERE IS C4A4.
CC   -!- POLYMORPHISM: THE C4A ALLELES CARRY THE BLOOD GROUP RODGERS WHILE
CC       THE C4B ALLELES CARRY THE BLOOD GROUP CHIDO.
CC   -!- DISEASE: Defects in C4A are the cause of C4A deficiency
CC       [MIM:120810].
CC   -!- DISEASE: The C4A6 allotype is totally deficient in hemolytic
CC       activity.
CC   -!- MISCELLANEOUS: C4A ALLOTYPES REACT MORE RAPIDLY WITH THE AMINO
CC       GROUP OF PEPTIDE ANTIGENS WHILE C4B ALLOTYPES REACT MORE RAPIDLY
CC       WITH THE HYDROXYL GROUP OF CARBOHYDRATE ANTIGENS.
CC   -!- MISCELLANEOUS: C4 IS A MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-III
CC       PROTEIN.
CC   -!- SIMILARITY: TO C3, C5 AND ALPHA-2-MACROGLOBULIN.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02403; AAB59537.1; -.
DR   EMBL; K02404; AAA59651.1; -.
DR   EMBL; M59815; AAA51855.1; -.
DR   EMBL; M59816; AAA51855.1; JOINED.
```

```
DR   EMBL; U24578; AAA99717.1; -.
DR   EMBL; AF019413; AAB67980.1; -.
DR   EMBL; AL049547; CAB89302.1; -.
DR   EMBL; M14823; AAA35617.1; -.
DR   EMBL; M14824; AAA52292.1; -.
DR   EMBL; X77491; CAA54627.1; -.
DR   PIR; I56095; C4HU.
DR   HSSP; P01031; 1KJS.
DR   SWISS-2DPAGE; P01028; HUMAN.
DR   Genew; HGNC:1323; C4A.
DR   Genew; HGNC:1324; C4B.
DR   MIM; 120810; -:
DR   MIM; 120820; -.
DR   MIM; 120790; -.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0006958; P:complement activation, classical pathway; NAS.
DR   GO; GO:0006954; P:inflammatory response; NAS.
DR   GO; GO:0006937; P:regulation of muscle contraction; NAS.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Plasma; Glycoprotein; Sulfation; Signal;
KW   Inflammatory response; Polymorphism; Disease mutation;
KW   Blood group antigen; Thioester bond.
FT   SIGNAL        1     19
FT   CHAIN        20    675       COMPLEMENT C4, BETA CHAIN.
FT   PROPEP      676    679
FT   CHAIN       680   1446       COMPLEMENT C4, ALPHA CHAIN.
FT   PROPEP     1447   1453
FT   CHAIN      1454   1744       COMPLEMENT C4, GAMMA CHAIN.
FT   PEPTIDE     680    756       C4A ANAPHYLATOXIN.
FT   DOMAIN      702    736       ANAPHYLATOXIN-LIKE.
FT   DISULFID    702    728       BY SIMILARITY.
FT   DISULFID    703    735       BY SIMILARITY.
FT   DISULFID    716    736       BY SIMILARITY.
FT   CROSSLNK   1010   1013       Isoglutamyl cysteine thioester (Cys-Gln).
FT   MOD_RES    1417   1417       SULFATION.
FT   MOD_RES    1420   1420       SULFATION.
FT   MOD_RES    1422   1422       SULFATION.
FT   CARBOHYD    226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    862    862       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1328   1328       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1391   1391       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     477    477       R -> W (in allotype C4A6).

  Query Match             45.2%;  Score 51.5;  DB 1;  Length 1744;
  Best Local Similarity   52.4%;  Pred. No. 4.4;
  Matches   11;  Conservative    3;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 CCYDGAS-VNNDETCEQRAAR 20
              || || : :    :|||||||
Db        702 CCQDGVTRLPMMRSCEQRAAR 722


RESULT 10
Y32K_BNYVG
ID   Y32K_BNYVG     STANDARD;      PRT;   282 AA.
AC   P19231;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   RNA-4 hypothetical 31.9 kDa protein.
OS   Beet necrotic yellow vein mosaic virus (isolate G1).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
OX   NCBI_TaxID=12257;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bouzoubaa S., Guilley H., Jonard G., Richards K., Putz C.;
RT   "Nucleotide sequence analysis of RNA-3 and RNA-4 of beet necrotic
RT   yellow vein virus, isolates P2 and G1.";
RL   J. Gen. Virol. 66:1553-1564(1985).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M36897; AAA42800.1; -.
DR   PIR; C44503; C44503.
DR   InterPro; IPR007004; DUF656.
DR   Pfam; PF04920; DUF656; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   282 AA;  31869 MW;  AA7C0351C54FE0CC CRC64;

  Query Match             42.1%;  Score 48;  DB 1;  Length 282;
  Best Local Similarity   57.1%;  Pred. No. 2.2;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNNDETCE 15
              || |  | :|| ||
Db        115 CYSGVDVLSDELCE 128


RESULT 11
OVOS_CHICK
ID   OVOS_CHICK     STANDARD;      PRT;  1473 AA.
AC   P20740;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Ovostatin precursor (Ovomacroglobulin).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oviduct;
RX   MEDLINE=95218210; PubMed=7535598;
RA   Nielsen K.L., Sottrup-Jensen L., Nagase H., Thoegersen H.C.,
RA   Etzerodt M.;
RT   "Amino acid sequence of hen ovomacroglobulin (ovostatin) deduced from
RT   cloned cDNA.";
RL   DNA Seq. 5:111-119(1994).
RN   [2]
RP   SEQUENCE OF 37-49.
RC   TISSUE=Egg white;
RX   MEDLINE=83238315; PubMed=6408074;
RA   Nagase H., Harris E.D. Jr., Woessner J.F., Brew K.;
RT   "Ovostatin: a novel proteinase inhibitor from chicken egg white. I.
RT   Purification, physicochemical properties, and tissue distribution of
RT   ovostatin.";
RL   J. Biol. Chem. 258:7481-7489(1983).
RN   [3]
RP   SEQUENCE OF 976-1028.
RC   TISSUE=Egg white;
RX   MEDLINE=93192299; PubMed=7680577;
RA   Nielsen K.L., Sottrup-Jensen L.;
RT   "Evidence from sequence analysis that hen egg-white ovomacroglobulin
RT   (ovostatin) is devoid of an internal beta-Cys-gamma-Glu thiol
RT   ester.";
```

```
RL   Biochim. Biophys. Acta 1162:230-232(1993).
CC   -!- FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A
CC       UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH,
CC       CALLED THE 'BAIT REGION' WHICH CONTAINS SPECIFIC CLEAVAGE SITES
CC       FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT
CC       REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH
CC       TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST
CC       LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR
CC       WEIGHT SUBSTRATES IS GREATLY REDUCED).
CC   -!- SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULFIDE-
CC       LINKED CHAINS.
CC   -!- PTM: THIS PROTEIN LACKS THE THIOESTER BOND FOUND IN OTHER MEMBERS
CC       OF THIS FAMILY.
CC   -!- PTM: CONTAINS 56 MOL GLUCOSAMINE PER MOL SUBUNIT.
CC   -!- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
CC       INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X78801; CAA55384.1; -.
DR   EMBL; X78801; CAA55385.1; ALT_INIT.
DR   PIR; I50671; A20872.
DR   HSSP; P01023; 1BV8.
DR   GO; GO:0017114; F:wide-spectrum protease inhibitor activity; NAS.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; MacroglobInA2.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; FALSE_NEG.
KW   Serine protease inhibitor; Glycoprotein; Bait region; Signal.
FT   SIGNAL        1     36
FT   CHAIN        37   1473       OVOSTATIN.
FT   CARBOHYD     67     67       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     89     89       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    191    191       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    342    342       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    403    403       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    527    527       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    588    588       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    757    757       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1141   1141       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1221   1221       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1315   1315       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1347   1347       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     46     46       M -> V (IN REF. 2).
SQ   SEQUENCE   1473 AA;  166354 MW;  A33C6847A14179BF CRC64;

  Query Match             40.4%;  Score 46;  DB 1;  Length 1473;
  Best Local Similarity   53.3%;  Pred. No. 27;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          6 ASVNNDETCEQRAAR 20
              |  |:|| ||: | |
Db        883 AETNDDEACEEEALR 897


RESULT 12
Y125_MYCPN
ID   Y125_MYCPN     STANDARD;      PRT;   281 AA.
AC   P75511;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein MG125 homolog (A65_orf281).
GN   MPN264 OR MP569.
OS   Mycoplasma pneumoniae.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2104;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE=97105885; PubMed=8948633;
RA   Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA   Herrmann R.;
RT   "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT   pneumoniae.";
RL   Nucleic Acids Res. 24:4420-4449(1996).
CC   -!- SIMILARITY: BELONGS TO THE COF/YBHA/YIDA/YIGL (E.COLI) / YCSE/YXEH
CC       (B.SUBTILIS) FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000056; AAB96217.1; -.
DR   PIR; S73895; S73895.
DR   InterPro; IPR006379; HAD_SF_IIB.
DR   InterPro; IPR005834; Hydrolase.
DR   InterPro; IPR000150; Hypothet_cof.
DR   Pfam; PF00702; Hydrolase; 1.
DR   TIGRFAMs; TIGR00099; Cof-subfamily; 1.
DR   TIGRFAMs; TIGR01484; HAD-SF-IIB; 1.
DR   PROSITE; PS01228; COF_1; 1.
DR   PROSITE; PS01229; COF_2; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   281 AA;  32614 MW;  BF44564E7C7FBF11 CRC64;

  Query Match             39.0%;  Score 44.5;  DB 1;  Length 281;
  Best Local Similarity   47.4%;  Pred. No. 8;
  Matches    9;  Conservative    3;  Mismatches    4;  Indels    3;  Gaps    1;

Qy          1 CCYDGA---SVNNDETCEQ 16
              ||| ||    :||::  ||
Db         71 CCYGGAKLYQLNNNQPQEQ 89


RESULT 13
CFAH_BOVIN
ID   CFAH_BOVIN     STANDARD;      PRT;   685 AA.
AC   Q28085;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Complement factor H (H factor 1) (Fragments).
GN   HF1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE OF 17-685 FROM N.A., AND SEQUENCE OF 1-16.
RC   TISSUE=Liver;
RX   MEDLINE=96202005; PubMed=8615824;
RA   Soames C.J., Day A.J., Sim R.B.;
RT   "Prediction from sequence comparisons of residues of factor H involved
RT   in the interaction with complement component C3b.";
RL   Biochem. J. 315:523-531(1996).
CC   -!- FUNCTION: Factor H functions as a cofactor in the inactivation of
CC       C3b by factor I and also increases the rate of dissociation of the
CC       C3bBb complex (C3 convertase) and the (C3b)NBB complex (C5
CC       convertase) in the alternative complement pathway (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Synthesized by liver and secreted in plasma.
CC   -!- SIMILARITY: Contains at least 13 Sushi (SCR) domains.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X98697; CAA67257.1; -.
DR   PIR; S65551; S65551.
DR   HSSP; P10998; 1VVD.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00084; sushi; 11.
DR   SMART; SM00032; CCP; 11.
KW   Complement alternate pathway; Plasma; Repeat; Sushi.
FT   NON_TER       1      1
FT   NON_CONS     16     17
FT   DOMAIN      <17     67       SUSHI 2.
FT   DOMAIN       70    131       SUSHI 3.
FT   DOMAIN      134    188       SUSHI 4.
FT   DOMAIN      191    246       SUSHI 5.
FT   DOMAIN      249    309       SUSHI 6.
FT   DOMAIN      311    366       SUSHI 7.
FT   DOMAIN      370    429       SUSHI 8.
FT   DOMAIN      431    488       SUSHI 9.
FT   DOMAIN      491    547       SUSHI 10.
FT   DOMAIN      553    609       SUSHI 11.
FT   DOMAIN      614    668       SUSHI 12.
FT   DOMAIN      675   >685       SUSHI 13.
FT   DISULFID     39     66       BY SIMILARITY.
FT   DISULFID     71    117       BY SIMILARITY.
FT   DISULFID    103    130       BY SIMILARITY.
FT   DISULFID    135    176       BY SIMILARITY.
FT   DISULFID    162    187       BY SIMILARITY.
FT   DISULFID    192    234       BY SIMILARITY.
FT   DISULFID    219    245       BY SIMILARITY.
FT   DISULFID    250    297       BY SIMILARITY.
FT   DISULFID    280    308       BY SIMILARITY.
FT   DISULFID    312    354       BY SIMILARITY.
FT   DISULFID    339    365       BY SIMILARITY.
FT   DISULFID    371    417       BY SIMILARITY.
FT   DISULFID    400    428       BY SIMILARITY.
FT   DISULFID    432    476       BY SIMILARITY.
FT   DISULFID    459    487       BY SIMILARITY.
FT   DISULFID    492    534       BY SIMILARITY.
FT   DISULFID    520    546       BY SIMILARITY.
FT   DISULFID    554    597       BY SIMILARITY.
FT   DISULFID    583    608       BY SIMILARITY.
FT   DISULFID    615    656       BY SIMILARITY.
FT   DISULFID    642    667       BY SIMILARITY.
FT   NON_TER     685    685
SQ   SEQUENCE   685 AA;  77536 MW;  69FC9DC8D530E872 CRC64;

  Query Match             39.0%;  Score 44.5;  DB 1;  Length 685;
  Best Local Similarity   44.4%;  Pred. No. 21;
  Matches    8;  Conservative    4;  Mismatches    3;  Indels    3;  Gaps    1;

Qy          2 CYDGASVNNDE---TCEQ 16
              ||:| |: ||:   || :
Db        339 CYEGYSLQNDQNTMTCTE 356


RESULT 14
UN36_CAEEL
ID   UN36_CAEEL     STANDARD;      PRT;  1249 AA.
AC   P34374; P34372; P34373;
DT   01-FEB-1994 (Rel. 28, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Voltage-dependent calcium channel unc-36 precursor (Uncoordinated
DE   protein 36).
GN   UNC-36 OR UNC-72 OR C50C3.9/C50C3.10/C50C3.11.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   REVISIONS.
RA   Waterston R.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   FUNCTION, AND TISSUE SPECIFICITY.
RX   PubMed=8807295;
RA   Schafer W.R., Sanchez B.M., Kenyon C.J.;
RT   "Genes affecting sensitivity to serotonin in Caenorhabditis elegans.";
RL   Genetics 143:1219-1230(1996).
CC   -!- FUNCTION: May act as an auxiliary subunit of the UNC-2 voltage-
CC       gated calcium channel which appears to trigger calcium-activated
CC       signaling pathways that control the serotonin response. Inhibiting
CC       serotonin sensitivity of the vulval muscles results in egg laying
CC       defects. May act in both neurons and muscle cells to enhance motor
CC       activity.
CC   -!- TISSUE SPECIFICITY: Decendants of the cells AB and AB.p (that give
CC       rise to nearly all nonpharyngeal neuorns), decendants of P1 (that
CC       give rise to body muscle) and cell lineages that give rise to the
CC       adult and juvenile motor neurons.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L14433; AAA27971.3; -.
DR   PIR; S44617; S44617.
DR   WormPep; C50C3.9; CE32168.
DR   InterPro; IPR004010; Cache.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF02743; Cache; 1.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Behavior; Ionic channel; Ion transport; Voltage-gated channel;
KW   Calcium channel; Signal; Glycoprotein.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20   1249       VOLTAGE-DEPENDENT CALCIUM CHANNEL UNC-36.
FT   DOMAIN      250    479       VWFA.
FT   CARBOHYD    100    100       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    140    140       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    146    146       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    302    302       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    520    520       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    558    558       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    757    757       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD    838    838       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    903    903       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    923    923       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1130   1130       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1249 AA;  144375 MW;  533FF6059D375FB2 CRC64;

  Query Match             39.0%;  Score 44.5;  DB 1;  Length 1249;
  Best Local Similarity   29.4%;  Pred. No. 40;
  Matches   10;  Conservative    3;  Mismatches    6;  Indels   15;  Gaps    1;

Qy          2 CYDGA--------------SVNNDETCEQRAAR 20
              ||| :                | |:||||:   |
Db       1170 CYDESECSMEISNQVPFGFEEVKNEETCEENEKR 1203


RESULT 15
ARLY_VIBPA
ID   ARLY_VIBPA     STANDARD;      PRT;   624 AA.
AC   P59620;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Bifunctional protein argH [Includes: Argininosuccinate lyase
DE   (EC 4.3.2.1) (Arginosuccinase) (ASAL); Probable acetyltransferase
DE   (EC 2.3.1.-)].
GN   ARGH OR VP2756.
OS   Vibrio parahaemolyticus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=670;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RIMD 2210633 / Serotype O3:K6;
RX   MEDLINE=22508454; PubMed=12620739;
RA   Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA   Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA   Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT   "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT   distinct from that of V. cholerae.";
RL   Lancet 361:743-749(2003).
CC   -!- CATALYTIC ACTIVITY: N-(L-arginino)succinate = fumarate + L-
CC       arginine.
CC   -!- PATHWAY: Arginine biosynthesis; eighth (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: In the N-terminal section; belongs to the lyase 1
CC       family. Argininosuccinate lyase subfamily.
CC   -!- SIMILARITY: In the C-terminal section; belongs to the
CC       acetyltransferase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP005082; BAC61019.1; -.
DR   HAMAP; MF_00006; fused; 1.
DR   PROSITE; PS00163; FUMARATE_LYASES; 1.
KW   Multifunctional enzyme; Arginine biosynthesis; Lyase; Transferase;
KW   Acyltransferase; Complete proteome.
FT   DOMAIN        1    466       ARGININOSUCCINATE LYASE.
FT   DOMAIN      467    624       PROBABLE ACETYLTRANSFERASE.
SQ   SEQUENCE   624 AA;  69357 MW;  DA5D4C815438ABCF CRC64;

  Query Match             38.6%;  Score 44;  DB 1;  Length 624;
  Best Local Similarity   50.0%;  Pred. No. 23;
  Matches    7;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNNDETCE 15
              |:||  || : | |
Db        345 CFDGIKVNGERTLE 358
```

Search completed: December 9, 2003, 14:07:51
Job time : 12 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 14:05:48 ; Search time 21 Seconds
(without alignments)
91.589 Million cell updates/sec

Title: US-09-651-685A-5
Perfect score: 114
Sequence: 1 CCYDGASVNNDETCEQRAAR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	109	95.6	1676	1	C5HU	complement C5 prec
2	94	82.5	74	2	A25408	complement C5 - bo
3	93	81.6	74	2	A01268	complement C5 - pi
4	84	73.7	77	2	A57689	complement C5a - r
5	84	73.7	1680	1	C5MS	complement C5 prec
6	56.5	49.6	1735	2	S54784	sex-limited protei
7	56.5	49.6	1736	2	A29176	sex-limited protei
8	56.5	49.6	1738	1	A24558	complement C4 prec
9	55.5	48.7	76	2	JL0036	complement C4a ana
10	55.5	48.7	77	2	A01265	complement C4 - bo
11	51.5	45.2	1744	1	C4HU	complement C4A pre
12	48	42.1	135	2	D44503	p19 protein - beet
13	48	42.1	282	2	JC5677	RNA4 protein - Bee
14	48	42.1	282	2	C44503	p31 protein - beet
15	47	41.2	834	2	JQ1965	hypothetical 94K p
16	46	40.4	1473	2	A20872	ovostatin precurso
17	45.5	39.9	1617	2	T28153	complement C4 - ch
18	45	39.5	159	2	E86760	conserved hypothet
19	45	39.5	394	2	T21013	hypothetical prote
20	44.5	39.0	161	2	S44619	C50C3.9 protein -
21	44.5	39.0	281	2	S73895	hypothetical prote
22	44.5	39.0	669	2	S65551	factor H - bovine
23	44	38.6	68	2	AE3441	17K common-antigen
24	44	38.6	2095	2	S29529	genome polyprotein
25	43.5	38.2	241	2	A59262	tetraspan TSPAN-1
26	43.5	38.2	1651	1	C3NJ	complement C3 prec
27	43	37.7	198	2	H90111	26S proteasome, be
28	43	37.7	337	2	T18654	hypothetical prote
29	43	37.7	461	2	AH0373	cysteine-tRNA liga
30	43	37.7	1645	2	AG1897	two-component hybr
31	43	37.7	2543	2	T31687	suface antigen - P
32	42.5	37.3	685	2	S78040	fibulin, splice fo
33	42.5	37.3	705	2	S34968	fibulin, splice fo
34	42.5	37.3	1620	2	I51339	complement compone
35	42.5	37.3	1673	2	I50806	complement compone
36	42	36.8	368	2	H75412	spermidine/putresc
37	42	36.8	1068	2	S01519	hypothetical prote
38	42	36.8	2092	2	S30026	genome polyprotein
39	42	36.8	2149	2	S18676	genome polyprotein
40	41.5	36.4	312	2	T32379	hypothetical prote
41	41.5	36.4	616	2	T32131	hypothetical prote
42	41.5	36.4	1101	2	T16840	hypothetical prote
43	41	36.0	187	2	T48093	hypothetical prote
44	41	36.0	208	2	C96833	hypothetical prote
45	41	36.0	217	2	A87990	protein W05H12.1 [

ALIGNMENTS

RESULT 1
C5HU
complement C5 precursor [validated] - human
N;Contains: C5a anaphylatoxin; C5b
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 08-Dec-2000
C;Accession: A40075; A27689; A01267; A01266; S15121
R;Haviland, D.L.; Haviland, J.C.; Fleischer, D.T.; Hunt, A.; Wetsel, R.A.
J. Immunol. 146, 362-368, 1991
A;Title: Complete cDNA sequence of human complement pro-C5. Evidence of truncated transc
A;Reference number: A40075; MUID:91079575; PMID:1984448
A;Accession: A40075
A;Molecule type: mRNA
A;Residues: 1-1676 <HAV>
A;Cross-references: GB:M57729; NID:g179982; PIDN:AAA51925.1; PID:g179983
A;Note: 518-Ser was also found
R;Wetsel, R.A.; Lemons, R.S.; Le Beau, M.M.; Barnum, S.R.; Noack, D.; Tack, B.F.
Biochemistry 27, 1474-1482, 1988
A;Title: Molecular analysis of human complement component C5: localization of the struct
A;Reference number: A27689; MUID:88209511; PMID:3365401
A;Accession: A27689
A;Molecule type: mRNA
A;Residues: 412-1676 <WET>
A;Cross-references: GB:M65134; GB:M18879; NID:g179691; PIDN:AAA51856.1; PID:g179692
R;Fernandez, H.N.; Hugli, T.E.
J. Biol. Chem. 253, 6955-6964, 1978
A;Title: Primary structural analysis of the polypeptide portion of human C5a anaphylatox
A;Reference number: A01267; MUID:79005687; PMID:690134
A;Accession: A01267
A;Molecule type: protein
A;Residues: 678-751 <FER>
R;Lundwall, A.B.; Wetsel, R.A.; Kristensen, T.; Whitehead, A.S.; Woods, D.E.; Ogden, R.C
J. Biol. Chem. 260, 2108-2112, 1985
A;Title: Isolation and sequence analysis of a cDNA clone encoding the fifth complement c
A;Reference number: A01266; MUID:85130937; PMID:2579066
A;Accession: A01266
A;Molecule type: mRNA
A;Residues: 412-854,'SLALSPRLECNGKISGHCKLRLPGSSDSPASASQVAGITGTHHHAQPT' <LUN>
A;Cross-references: GB:K02874
A;Note: the carboxyl-terminal part of the sequence in this report appears to be derived
R;Bohnsack, J.F.; Mollison, K.W.; Buko, A.M.; Ashworth, J.C.; Hill, H.R.
Biochem. J. 273, 635-640, 1991
A;Title: Group B streptococci inactivate complement component C5a by enzymic cleavage at
A;Reference number: S15121; MUID:91144547; PMID:1996961
A;Contents: annotation
C;Comment: Complement C5 contains two disulfide-linked chains, formed by removal of four
(beta and alpha' chains).
C;Comment: Activation of C5 initiates the spontaneous assembly of the late complement co
is the foundation upon which the membrane attack complex is assembled.
C;Comment: C5a has potent spasmogenic and chemotactic activity.
C;Genetics:
A;Gene: GDB:C5

```
A;Cross-references: GDB:119734; OMIM:120900
A;Map position: 9q34.1-9q34.1
C;Superfamily: alpha-2-macroglobulin
C;Keywords: complement alternate pathway; complement pathway; cytolysis; glycoprotein; i
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-673,678-1676/Product: complement C5 #status predicted <MAT>
F;19-673,752-1676/Product: C5b #status predicted <C5B>
F;19-673/Product: complement C5 and C5b beta chain #status predicted <C5BB>
F;678-1676/Product: complement C5 alpha chain #status predicted <C5A>
F;678-751/Product: C5a anaphylatoxin #status experimental <C5T>
F;752-1676/Product: C5b alpha' chain #status predicted <C5BA>
F;567-810,634-669,698-724,699-731,711-732,866-1527,1101-1159,1375-1505,1405-1474,1520-15
F;741/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;751-752/Cleavage site: Arg-Leu (C5 convertase) #status experimental
F;911,1115,1630/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             95.6%;  Score 109;  DB 1;  Length 1676;
  Best Local Similarity   95.0%;  Pred. No. 1.6e-08;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||||||||||||
Db        698 CCYDGACVNNDETCEQRAAR 717


RESULT 2
A25408
complement C5 - bovine (fragment)
N;Contains: C5a anaphylatoxin; C5b
C;Species: Bos primigenius taurus (cattle)
C;Date: 02-Jun-1988 #sequence_revision 15-Oct-1994 #text_change 16-Feb-1997
C;Accession: A25408
R;Gennaro, R.; Simonic, T.; Negri, A.; Mottola, C.; Secchi, C.; Ronchi, S.; Romeo, D.
Eur. J. Biochem. 155, 77-86, 1986
A;Title: C5a fragment of bovine complement. Purification, bioassays, amino-acid sequence
A;Reference number: A25408; MUID:86136134; PMID:3081348
A;Accession: A25408
A;Molecule type: protein
A;Residues: 1-74 <GEN>
C;Comment: Complement C5 contains two disulfide-linked chains, formed by removal of four
 (beta and alpha' chains).
C;Comment: Activation of C5 initiates the spontaneous assembly of the late complement co
is the foundation upon which the membrane attack complex is assembled.
C;Comment: C5a has potent spasmogenic and chemotactic activity.
C;Superfamily: alpha-2-macroglobulin
C;Keywords: complement alternate pathway; complement pathway; cytolysis; glycoprotein; i
F;1-74/Product: C5a anaphylatoxin #status experimental <C5T>
F;21-47,22-54,34-55/Disulfide bonds: #status predicted

  Query Match             82.5%;  Score 94;  DB 2;  Length 74;
  Best Local Similarity   80.0%;  Pred. No. 1.7e-07;
  Matches   16;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||  |:|||||:||||
Db         21 CCYDGAHRNDDETCEERAAR 40


RESULT 3
A01268
complement C5 - pig (fragment)
N;Contains: C5a anaphylatoxin; C5b
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 29-Jul-1981 #sequence_revision 15-Oct-1994 #text_change 16-Feb-1997
C;Accession: A01268; A26248
R;Gerard, C.; Hugli, T.E.
J. Biol. Chem. 255, 4710-4715, 1980
A;Title: Amino acid sequence of the anaphylatoxin from the fifth component of porcine co
A;Reference number: A01268; MUID:80182137; PMID:7372604
A;Accession: A01268
A;Molecule type: protein
A;Residues: 1-74 <GER>
R;Gerard, C.; Hugli, T.E.
Proc. Natl. Acad. Sci. U.S.A. 78, 1833-1837, 1981
A;Reference number: A26247; MUID:81199549; PMID:6940191
A;Contents: annotation; active region
A;Note: although Arg-74 is not essential, residues 72-74 (Leu-Gly-Arg) are required for
R;Zimmermann, B.; Vogt, W.
Hoppe-Seyler's Z. Physiol. Chem. 365, 151-158, 1984
A;Reference number: A26248; MUID:84184201; PMID:6714942
A;Contents: disulfide bonds
A;Accession: A26248
A;Molecule type: protein
A;Residues: 1-64,'E',66-73 <ZIM>
C;Comment: Complement C5 contains two disulfide-linked chains, formed by removal of four
 (beta and alpha' chains).
C;Comment: Activation of C5 initiates the spontaneous assembly of the late complement co
is the foundation upon which the membrane attack complex is assembled.
C;Comment: C5a has potent spasmogenic and chemotactic activity.
C;Superfamily: alpha-2-macroglobulin
C;Keywords: complement alternate pathway; complement pathway; cytolysis; glycoprotein; i
F;1-74/Product: C5a anaphylatoxin #status experimental <C5T>
F;21-47,22-54,34-55/Disulfide bonds: #status experimental

  Query Match             81.6%;  Score 93;  DB 2;  Length 74;
  Best Local Similarity   80.0%;  Pred. No. 2.4e-07;
  Matches   16;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||  |:|||||:||||
Db         21 CCYDGAYRNDDETCEERAAR 40


RESULT 4
A57689
complement C5a - rat (fragment)
N;Contains: C5a anaphylatoxin; C5b
C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Feb-1996 #sequence_revision 31-Jan-1997 #text_change 16-Feb-1997
C;Accession: A57689
R;Cui, L.; Carney, D.F.; Hugli, T.E.
Protein Sci. 3, 1169-1177, 1994
A;Title: Primary structure and functional characterization of rat C5a: an anaphylatoxin
A;Reference number: A57689; MUID:95078724; PMID:7987212
A;Accession: A57689
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-77 <CUI>
C;Complex: Complement C5 contains two disulfide-linked chains, formed by removal of four
 (beta and alpha' chains).
C;Function:
A;Description: activation of C5 initiates spontaneous assembly of the late complement co
is the foundation for assembly of the membrane attack complex
A;Pathway: complement alternate pathway; complement pathway
A;Note: C5a has potent spasmogenic and chemotactic activity
C;Superfamily: alpha-2-macroglobulin
C;Keywords: complement alternate pathway; complement pathway; cytolysis; glycoprotein; i
F;1-77/Product: C5a anaphylatoxin #status experimental <C5T>
F;24-50,25-57,37-58/Disulfide bonds: #status predicted

  Query Match             73.7%;  Score 84;  DB 2;  Length 77;
  Best Local Similarity   75.0%;  Pred. No. 6e-06;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||||||  |  |||||| ||
Db         24 CCYDGARENKYETCEQRVAR 43


RESULT 5
C5MS
complement C5 precursor - mouse
N;Contains: C5a anaphylatoxin; C5b
C;Species: Mus musculus (house mouse)
C;Date: 19-Nov-1988 #sequence_revision 15-Oct-1994 #text_change 18-Jun-1999
C;Accession: A35530; A27538; A40429
```

R;Wetsel, R.A.; Fleischer, D.T.; Haviland, D.L.
J. Biol. Chem. 265, 2435-2440, 1990
A;Title: Deficiency of the murine fifth complement component (C5). A 2-base pair gene de
A;Reference number: A35530; MUID:90153853; PMID:2303408
A;Accession: A35530
A;Molecule type: mRNA
A;Residues: 1-215,'L' <WET>
A;Cross-references: GB:M35526; GB:J05234; NID:g192302; PIDN:AAA37348.1; PID:g309123
R;Wetsel, R.A.; Ogata, R.T.; Tack, B.F.
Biochemistry 26, 737-743, 1987
A;Title: Primary structure of the fifth component of murine complement.
A;Reference number: A27538; MUID:87185363; PMID:2436653
A;Accession: A27538
A;Molecule type: mRNA
A;Residues: 'PGL',44-1680 <WET2>
R;Haviland, D.L.; Haviland, J.C.; Fleischer, D.T.; Wetsel, R.A.
J. Biol. Chem. 266, 11818-11825, 1991
A;Title: Structure of the murine fifth complement component (C5) gene. A large, highly i
nt component genes.
A;Reference number: A40429; MUID:91268053; PMID:1711041
A;Accession: A40429
A;Molecule type: DNA
A;Residues: 1-15 <HAV>
A;Cross-references: GB:M64852
C;Comment: Complement C5 contains two disulfide-linked chains, formed by removal of four
(beta and alpha' chains).
C;Comment: Activation of C5 initiates the spontaneous assembly of the late complement co
is the foundation upon which the membrane attack complex is assembled.
C;Comment: C5a has potent spasmogenic and chemotactic activity.
C;Genetics:
A;Map position: 2
A;Introns: 22/3; 86/3; 140/3; 164/3; 195/2; 223/1; 253/2; 291/3; 334/1; 372/3; 434/3; 50
3; 1224/1; 1292/3; 1343/3; 1364/3; 1392/1; 1411/2; 1445/3; 1470/3; 1506/1; 1534/1; 1564/
C;Superfamily: alpha-2-macroglobulin
C;Keywords: complement alternate pathway; complement pathway; cytolysis; glycoprotein; i
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-674,679-1679/Product: complement C5 #status predicted <MAT>
F;19-674,756-1679/Product: C5b #status predicted <C5B>
F;19-674/Product: complement C5 and C5b beta chain #status predicted <C5BB>
F;679-1679/Product: complement C5 alpha chain #status predicted <C5A>
F;679-755/Product: C5a anaphylatoxin #status predicted <C5T>
F;756-1679/Product: C5b alpha' chain #status predicted <C5BA>
F;567-814,635-670,702-728,703-735,715-736,870-1531,1105-1163,1379-1509,1409-1478,1524-15
F;915,1119,1633/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 73.7%; Score 84; DB 1; Length 1680;
Best Local Similarity 75.0%; Pred. No. 0.00011;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||| ||  ||||:| ||
Db        702 CCYDGARVNFYETCEERVAR 721
```

RESULT 6
S54784
sex-limited protein Slp(w7) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 16-Jul-1999
C;Accession: S54784; I48770
R;Ogata, R.T.; Zepf, N.E.
J. Immunol. 147, 2756-2763, 1991
A;Title: The murine Slp gene. Additional evidence that sex-limited protein has no biolog
A;Reference number: S54784; MUID:92013090; PMID:1918990
A;Accession: S54784
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-1735 <OGA>
A;Cross-references: EMBL:M64933; NID:g200988; PIDN:AAA40117.1; PID:g200989
R;Hemenway, C.; Kalff, M.; Stavenhagen, J.; Walthall, D.; Robins, D.
Nucleic Acids Res. 14, 2539-2554, 1986
A;Title: Sequence comparison of alleles of the fourth component of complement (C4) and s
A;Reference number: I48274; MUID:86176748; PMID:3008092
A;Accession: I48770
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 634-641,'L',643-828 <RES>
A;Cross-references: EMBL:X06455; NID:g54102; PIDN:CAA29761.1; PID:g899336
C;Genetics:
A;Introns: 22/2; 86/3; 154/1; 177/3; 207/2; 235/1; 267/2; 302/3; 347/1; 385/3; 444/3; 50
3; 1219/1; 1296/3; 1352/3; 1372/3; 1404/1; 1464/2; 1494/3; 1519/3; 1554/1; 1584/1; 1617/
C;Superfamily: alpha-2-macroglobulin

Query Match 49.6%; Score 56.5; DB 2; Length 1735;
Best Local Similarity 57.1%; Pred. No. 2;
Matches 12; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

```
Qy          1 CCYDG-ASVNNDETCEQRAAR 20
              || ||   :    ||||||||
Db        700 CCQDGMTKLPMKRTCEQRAAR 720
```

RESULT 7
A29176
sex-limited protein precursor - mouse
N;Alternate names: complement component C4-related sex-limited protein
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C;Accession: A29176; B21692; A41195; B29059; B60227; I79467
R;Ogata, R.T.; Sepich, D.S.
J. Immunol. 135, 4239-4244, 1985
A;Title: Murine sex-limited protein: complete cDNA sequence and comparison with murine f
A;Reference number: A29176; MUID:86060918; PMID:3840826
A;Accession: A29176
A;Molecule type: mRNA
A;Residues: 1-1736 <OGA>
R;Nonaka, M.; Takahashi, M.; Natsuume-Sakai, S.; Nonaka, M.; Tanaka, S.; Shimizu, A.; Ho
Proc. Natl. Acad. Sci. U.S.A. 81, 6822-6826, 1984
A;Title: Isolation of cDNA clones specifying the fourth component of mouse complement an
A;Reference number: A94013; MUID:85038607; PMID:6208559
A;Accession: B21692
A;Molecule type: mRNA
A;Residues: 651-749,'H',751-774,'D',776-802,921-1040 <NON>
R;Ogata, R.T.; Sepich, D.S.
Proc. Natl. Acad. Sci. U.S.A. 81, 4908-4911, 1984
A;Title: Genes for murine fourth complement component (C4) and sex-limited protein (Slp)
A;Reference number: A41195; MUID:84272739; PMID:6589636
A;Accession: A41195
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1345-1438,'R',1440-1544 <OG2>
A;Cross-references: GB:K02293; NID:g199611; PIDN:AAA39682.1; PID:g199612
R;Tosi, M.; Levi-Strauss, M.; Duponchel, C.; Meo, T.
Philos. Trans. R. Soc. Lond. 306, 389-394, 1984
A;Title: Sequence heterogeneity of murine complementary DNA clones related to the C4 and
A;Reference number: A93753
A;Accession: B29059
A;Molecule type: mRNA
A;Residues: 1255-1335,'G',1337-1373 <TOS>
R;Ogata, R.T.; Zepf, N.E.
Eur. J. Immunol. 20, 1607-1610, 1990
A;Title: C4 from C4-high and C4-low mouse strains have identical sequences in the region
A;Reference number: A60227; MUID:90353398; PMID:2387317
A;Accession: B60227
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1096-1139 <OG3>
A;Cross-references: GB:X55495
R;Nonaka, M.; Kimura, H.; Yeul, Y.D.; Yokoyama, S.; Nakayama, K.; Takahashi, M.
Proc. Natl. Acad. Sci. U.S.A. 83, 7883-7887, 1986
A;Title: Identification of the 5'-flanking regulatory region responsible for the differe
A;Reference number: I59084; MUID:87017050; PMID:3464002
A;Accession: I79467
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-21 <RES>

A;Cross-references: GB:M14226; NID:g199615; PIDN:AAA39684.1; PID:g554239
C;Genetics:
A;Gene: slp
C;Superfamily: alpha-2-macroglobulin
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1736/Product: sex-limited protein #status predicted <MAT>

Query Match 49.6%; Score 56.5; DB 2; Length 1736;
Best Local Similarity 57.1%; Pred. No. 2;
Matches 12; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

```
Qy          1 CCYDG-ASVNNDETCEQRAAR 20
              || ||   :    ||||||||
Db        700 CCQDGMTKLPMKRTCEQRAAR 720
```

RESULT 8
A24558
complement C4 precursor - mouse
N;Contains: classical-complement-pathway C3/C5 convertase (EC 3.4.21.43) C4b subunit; cc
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1989 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C;Accession: A24558; A25371; A21692; A30520; A60227; A22039; A29059; A01264; B41195; I59
R;Sepich, D.S.; Noonan, D.J.; Ogata, R.T.
Proc. Natl. Acad. Sci. U.S.A. 82, 5895-5899, 1985
A;Title: Complete cDNA sequence of the fourth component of murine complement.
A;Reference number: A24558; MUID:85298264; PMID:3862104
A;Accession: A24558
A;Molecule type: mRNA
A;Residues: 1-1738 <SEP>
A;Cross-references: GB:M11729; NID:g199080; PIDN:AAA39506.1; PID:g387420
A;Experimental source: strain B10.WR
R;Nonaka, M.; Nakayama, K.; Yeul, Y.D.; Takahashi, M.
J. Biol. Chem. 260, 10936-10943, 1985
A;Title: Complete nucleotide and derived amino acid sequences of the fourth component of
A;Reference number: A25371; MUID:85289294; PMID:2993295
A;Accession: A25371
A;Molecule type: mRNA
A;Residues: 1-131,'Y',133-326,'E',328-569,'E',571-1323,'N',1325-1441,'K',1443-1452,'V',1
A;Cross-references: GB:M11789
A;Experimental source: strain FM
R;Nonaka, M.; Takahashi, M.; Natsuume-Sakai, S.; Nonaka, M.; Tanaka, S.; Shimizu, A.; Ho
Proc. Natl. Acad. Sci. U.S.A. 81, 6822-6826, 1984
A;Title: Isolation of cDNA clones specifying the fourth component of mouse complement an
A;Reference number: A94013; MUID:85038607; PMID:6208559
A;Accession: A21692
A;Molecule type: mRNA
A;Residues: 651-719,'G',721-738,'AI',741-805 <NO2>
A;Cross-references: GB:M12970
A;Experimental source: strain FM
R;Taillon-Miller, P.A.; Shreffler, D.C.
J. Immunol. 141, 2382-2387, 1988
A;Title: Structural basis for the C4d.1/C4d.2 serologic allotypes of murine complement c
A;Reference number: A30520; MUID:89009745; PMID:2459207
A;Accession: A30520
A;Molecule type: DNA
A;Residues: 961-1205,'Q',1207-1290 <TAI>
A;Cross-references: GB:M23186; NID:g340821; PIDN:AAA40487.1; PID:g554392
A;Experimental source: strain B10.BR
R;Ogata, R.T.; Zepf, N.E.
Eur. J. Immunol. 20, 1607-1610, 1990
A;Title: C4 from C4-high and C4-low mouse strains have identical sequences in the region
A;Reference number: A60227; MUID:90353398; PMID:2387317
A;Accession: A60227
A;Molecule type: DNA
A;Residues: 1099-1142 <OGA>
A;Cross-references: GB:X55493; NID:g287729; PIDN:CAA39112.1; PID:g287730
R;Levi-Strauss, M.; Tosi, M.; Steinmetz, M.; Klein, J.; Meo, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 1746-1750, 1985
A;Title: Multiple duplications of complement C4 gene correlate with H-2-controlled testo
A;Reference number: A22039; MUID:85166208; PMID:3856857
A;Accession: A22039
A;Molecule type: mRNA
A;Residues: 1105-1118,'A',1120-1189,'T',1191-1449 <LEV>
A;Cross-references: GB:K02798; NID:g199281; PIDN:AAC42021.1; PID:g199282
A;Experimental source: strain B10.W7R
R;Tosi, M.; Levi-Strauss, M.; Duponchel, C.; Meo, T.
Philos. Trans. R. Soc. Lond. 306, 389-394, 1984
A;Title: Sequence heterogeneity of murine complementary DNA clones related to the C4 and
A;Reference number: A93753
A;Accession: A29059
A;Molecule type: mRNA
A;Residues: 1258-1376 <TOS>
A;Cross-references: GB:K02798
R;Ogata, R.T.; Shreffler, D.C.; Sepich, D.S.; Lilly, S.P.
Proc. Natl. Acad. Sci. U.S.A. 80, 5061-5065, 1983
A;Title: cDNA clone spanning the alpha-gamma subunit junction in the precursor of the mu
A;Reference number: A01264; MUID:83273751; PMID:6192448
A;Accession: A01264
A;Molecule type: mRNA
A;Residues: 1360-1400,'S',1402-1511 <OG2>
A;Cross-references: GB:K00019; NID:g199259; PIDN:AAA39554.1; PID:g554209
A;Experimental source: strain B10.W7R
R;Ogata, R.T.; Sepich, D.S.
Proc. Natl. Acad. Sci. U.S.A. 81, 4908-4911, 1984
A;Title: Genes for murine fourth complement component (C4) and sex-limited protein (Slp)
A;Reference number: A41195; MUID:84272739; PMID:6589636
A;Accession: B41195
A;Molecule type: mRNA
A;Residues: 1360-1400,'S',1402-1511 <OG3>
A;Cross-references: GB:K00019; NID:g199259; PIDN:AAA39554.1; PID:g554209
A;Experimental source: strain B10.W7R
R;Nonaka, M.; Kimura, H.; Yeul, Y.D.; Yokoyama, S.; Nakayama, K.; Takahashi, M.
Proc. Natl. Acad. Sci. U.S.A. 83, 7883-7887, 1986
A;Title: Identification of the 5'-flanking regulatory region responsible for the differe
A;Reference number: I59084; MUID:87017050; PMID:3464002
A;Accession: I59084
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-21 <RES>
A;Cross-references: GB:M14225; NID:g199291; PIDN:AAA39563.1; PID:g554211
R;Hemenway, C.; Kalff, M.; Stavenhagen, J.; Walthall, D.; Robins, D.
Nucleic Acids Res. 14, 2539-2554, 1986
A;Title: Sequence comparison of alleles of the fourth component of complement (C4) and s
A;Reference number: I48274; MUID:86176748; PMID:3008092
A;Accession: I48274
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 591-603,'M',605-1323,'N',1325-1452,'V',1454-1585,'Q',1587-1738 <RE2>
A;Cross-references: EMBL:X05314; NID:g50241; PIDN:CAA28936.1; PID:g50242
R;Nonaka, M.; Nakayama, K.; Yeul, Y.D.; Shimizu, A.; Takahashi, M.
Immunol. Rev. 87, 81-99, 1985
A;Title: Molecular cloning and characterization of complementary and genomic DNA clones
A;Reference number: I54567; MUID:86031969; PMID:2997024
A;Accession: I54567
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-128 <RE3>
A;Cross-references: GB:M12968; NID:g199267; PIDN:AAA39558.1; PID:g199270
A;Accession: I69023
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1724-1738 <RE4>
A;Cross-references: GB:M12969; NID:g199268; PIDN:AAA39559.1; PID:g387439
C;Comment: This protein is synthesized as a single-chain precursor and, prior to secreti
by disulfide bonds.
C;Comment: The activation of complement C4 by complement subcomponent C1s releases the C
of complement factor 2 to form the classical-complement-pathway C3 convertase. The C4b,
ay C5 convertase.
C;Comment: C4a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: The activity of C4b is regulated by proteolytic cleavage involving C4b-bindin
C;Genetics:
A;Introns: 22/2; 86/3; 977/3; 1047/3; 1073/3; 1125/3; 1164/3; 1221/3
A;Note: the list of introns is incomplete
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; complement classical pathway; glycoprotein; hydrolase; inflamma

F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-673/Product: complement C4 beta chain #status predicted <BET>
F;20-673,754-1440,1448-1738/Product: complement C4b #status predicted <C4B>
F;678-1440/Product: complement C4 alpha chain #status predicted <ALP>
F;678-753/Product: C4a anaphylatoxin #status predicted <C4A>
F;754-843/Region: C4b-binding protein binding
F;953-1332/Product: C4d fragment #status predicted <C4D>
F;1448-1738/Product: complement C4 gamma chain #status predicted <GAM>
F;224,743,1387/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;753-754/Cleavage site: Arg-Asn (complement subcomponent C1s) #status predicted
F;1006-1009/Cross-link: thiolester (Cys-Gln) #status predicted

Query Match 49.6%; Score 56.5; DB 1; Length 1738;
Best Local Similarity 57.1%; Pred. No. 2;
Matches 12; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

```
Qy          1 CCYDG-ASVNNDETCEQRAAR 20
              || ||   :    ||||||||
Db        700 CCQDGMTKLPMKRTCEQRAAR 720
```

RESULT 9
JL0036
complement C4a anaphylatoxin - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 15-Nov-1996
C;Accession: JL0036
R;Cui, L.; Ferreri, K.; Hugli, T.E.
Mol. Immunol. 25, 663-671, 1988
A;Title: Structural characterization of the C4a anaphylatoxin from rat.
A;Reference number: JL0036; MUID:88334568; PMID:3262196
A;Accession: JL0036
A;Molecule type: protein
A;Residues: 1-76 <CUI>
C;Superfamily: alpha-2-macroglobulin
C;Keywords: complement classical pathway; glycoprotein
F;66/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 48.7%; Score 55.5; DB 2; Length 76;
Best Local Similarity 57.1%; Pred. No. 0.15;
Matches 12; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

```
Qy          1 CCYDG-ASVNNDETCEQRAAR 20
              || ||   :    ||||||||
Db         23 CCQDGMTKLPMARTCEQRAAR 43
```

RESULT 10
A01265
complement C4 - bovine (fragment)
N;Contains: C4a anaphylatoxin; classical-complement-pathway C3/C5 convertase (EC 3.4.21.
C;Species: Bos primigenius taurus (cattle)
C;Date: 13-Jun-1983 #sequence_revision 15-Oct-1994 #text_change 16-Feb-1997
C;Accession: A01265
R;Smith, M.A.; Gerrie, L.M.; Dunbar, B.; Fothergill, J.E.
Biochem. J. 207, 253-260, 1982
A;Title: Primary structure of bovine complement activation fragment C4a, the third anaph
A;Reference number: A01265; MUID:83126436; PMID:6760852
A;Accession: A01265
A;Molecule type: protein
A;Residues: 1-77 <SMI>
C;Superfamily: alpha-2-macroglobulin
C;Keywords: complement pathway; glycoprotein; hydrolase; inflammation; plasma; serine pr
F;1-77/Product: C4a anaphylatoxin #status experimental <C4T>

Query Match 48.7%; Score 55.5; DB 2; Length 77;
Best Local Similarity 57.1%; Pred. No. 0.15;
Matches 12; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

```
Qy          1 CCYDGAS-VNNDETCEQRAAR 20
              || || : :    ||||||||
Db         23 CCQDGLTRLPMARTCEQRAAR 43
```

RESULT 11
C4HU
complement C4A precursor [validated] - human
N;Contains: classical-complement-pathway C3/C5 convertase (EC 3.4.21.43) C4b subunit; co
C;Species: Homo sapiens (man)
C;Date: 25-Feb-1985 #sequence_revision 23-Aug-1996 #text_change 08-Dec-2000
C;Accession: I56095; A29177; B29177; A90845; A19311; A92337; S12866; A17265; A32335; A27
R;Yu, C.Y.
J. Immunol. 146, 1057-1066, 1991
A;Title: The complete exon-intron structure of a human complement component C4A gene. DN
A;Reference number: I56095; MUID:91108039; PMID:1988494
A;Accession: I56095
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1744 <RES>
A;Cross-references: GB:M59815; NID:g179672; PIDN:AAA51855.1; PID:g179674
R;Belt, K.T.; Yu, C.Y.; Carroll, M.C.; Porter, R.R.
Immunogenetics 21, 173-180, 1985
A;Title: Polymorphism of human complement component C4.
A;Reference number: A29177; MUID:85156269; PMID:3838531
A;Accession: A29177
A;Molecule type: DNA
A;Residues: 1-22 <BEL>
A;Cross-references: GB:M14823
A;Accession: B29177
A;Molecule type: DNA
A;Residues: 1056-1225 <BE3>
A;Cross-references: GB:M14824; NID:g179675; PIDN:AAA52292.1; PID:g553210
R;Belt, K.T.; Carroll, M.C.; Porter, R.R.
Cell 36, 907-914, 1984
A;Title: The structural basis of the multiple forms of human complement component C4.
A;Reference number: A90845; MUID:84156544; PMID:6546707
A;Accession: A90845
A;Molecule type: mRNA
A;Residues: 20-346,'S',348-417,'A',419-725,'P',727-1200,'S',1202-1285,'S',1287-1418,1422
A;Cross-references: GB:K02403
R;Carroll, M.C.; Porter, R.R.
Proc. Natl. Acad. Sci. U.S.A. 80, 264-267, 1983
A;Title: Cloning of a human complement component C4 gene.
A;Reference number: A19311; MUID:83117835; PMID:6572000
A;Accession: A19311
A;Molecule type: mRNA
A;Residues: 1195-1285,'S',1287-1294 <CAR>
A;Cross-references: GB:V00502; GB:J00080; NID:g30010; PIDN:CAA23760.1; PID:g1335030
R;Moon, K.E.; Gorski, J.P.; Hugli, T.E.
J. Biol. Chem. 256, 8685-8692, 1981
A;Title: Complete primary structure of human C4a anaphylatoxin.
A;Reference number: A92337; MUID:81264286; PMID:6167582
A;Accession: A92337
A;Molecule type: protein
A;Residues: 680-725,'PN',728-756 <MOO>
R;Hessing, M.; van't Veer, C.; Hackeng, T.M.; Bouma, B.N.; Iwanaga, S.
FEBS Lett. 271, 131-136, 1990
A;Title: Importance of the alpha(3)-fragment of complement C4 for the binding with C4b-b
A;Reference number: S12866; MUID:91032049; PMID:1699796
A;Accession: S12866
A;Molecule type: protein
A;Residues: 757,'X',759-771;980-990 <HES>
R;Campbell, R.D.; Gagnon, J.; Porter, R.R.
Biochem. J. 199, 359-370, 1981
A;Title: Amino acid sequence around the thiol and reactive acyl groups of human compleme
A;Reference number: A17265; MUID:82182029; PMID:6978711
A;Accession: A17265
A;Molecule type: protein
A;Residues: 957-1012,'E',1014-1044 <CAM>
R;Chakravarti, D.N.; Campbell, R.D.; Porter, R.R.
Mol. Immunol. 24, 1187-1197, 1987
A;Title: The chemical structure of the C4d fragment of the human complement component C4
A;Reference number: A32335; MUID:88094444; PMID:3696167
A;Accession: A32335
A;Molecule type: protein
A;Residues: 957-1012,'E',1014-1108,'I',1110-1175,'S',1177-1270,'V',1272-1336 <CHA>

```
A;Note: 1073-Gly, 1120-Leu, 1121-Ser, 1124-Ile, 1125-His, 1207-Ala, 1210-Arg were also f
R;Chakravarti, D.N.; Campbell, R.D.; Gagnon, J.
FEBS Lett. 154, 387-390, 1983
A;Title: Amino acid sequence of a polymorphic segment from fragment C4d of human complem
A;Reference number: A27600; MUID:83158189; PMID:6832377
A;Accession: A27600
A;Molecule type: protein
A;Residues: 1199-1270,'V',1272-1299,'V',1301-1304 <CH2>
R;Whitehead, A.S.; Goldberger, G.; Woods, D.E.; Markham, A.F.; Colten, H.R.
Proc. Natl. Acad. Sci. U.S.A. 80, 5387-5391, 1983
A;Title: Use of a cDNA clone for the fourth component of human complement (C4) for analy
A;Reference number: I58991; MUID:83299979; PMID:6577433
A;Accession: I58991
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1448-1474 <RE2>
A;Cross-references: GB:K00830; NID:g187772; PIDN:AAA36229.1; PID:g187773
R;Sargent, C.A.; Anderson, M.J.; Hsieh, S.L.; Kendall, E.; Gomez-Escobar, N.; Campbell,
Hum. Mol. Genet. 3, 481-488, 1994
A;Title: Characterisation of the novel gene G11 lying adjacent to the complement C4A ger
A;Reference number: I37396; MUID:94282044; PMID:8012361
A;Accession: I37399
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-21 <RE3>
A;Cross-references: EMBL:X77491; NID:g453410; PIDN:CAA54627.1; PID:g453411
C;Comment: This protein is synthesized as a single-chain precursor and, prior to secreti
by disulfide bonds.
C;Comment: The activation of complement C4 by complement subcomponent C1s releases the C
 of complement factor 2 to form the classical-complement-pathway C3 convertase. The C4b,
ay C5 convertase.
C;Comment: C4a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: The activity of C4b is regulated by proteolytic cleavage involving C4b-bindin
C;Comment: Residues 1446 or 1449 may be the carboxyl end of the alpha chain.
C;Comment: There are at least two genes coding for C4, C4A and C4B. Each gene has many a
C;Genetics:
A;Gene: GDB:C4A
A;Cross-references: GDB:119732; OMIM:120810
A;Map position: 6p21.3-6p21.3
A;Introns: 22/2; 88/3; 156/1; 179/3; 209/2; 237/1; 269/2; 304/3; 349/1; 387/3; 447/3; 50
3; 1226/1; 1303/3; 1359/3; 1379/3; 1411/1; 1473/2; 1503/3; 1528/3; 1563/1; 1593/1; 1626/
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; complement classical pathway; glycoprotein; hydrolase; inflamma
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-675/Product: complement C4 beta chain #status predicted <BET>
F;20-675,757-1446,1454-1744/Product: C4b #status predicted <C4B>
F;680-1446/Product: complement C4 alpha chain #status predicted <ALP>
F;680-756/Product: C4a anaphylatoxin #status experimental <C4A>
F;757-845/Region: C4b-binding protein binding
F;957-1336/Product: C4d fragment #status experimental <C4D>
F;1454-1744/Product: complement C4 gamma chain #status predicted <GAM>
F;756-757/Cleavage site: Arg-Ala (complement subcomponent C1s) #status experimental
F;1010-1013/Cross-link: thiolester (Cys-Gln) #status experimental
F;1328/Binding site: carbohydrate (Asn) (covalent) #status experimental

  Query Match             45.2%;  Score 51.5;  DB 1;  Length 1744;
  Best Local Similarity   52.4%;  Pred. No. 12;
  Matches   11;  Conservative    3;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 CCYDGAS-VNNDETCEQRAAR 20
              || || : :    :|||||||
Db        702 CCQDGVTRLPMMRSCEQRAAR 722


RESULT 12
D44503
p19 protein - beet necrotic yellow vein mosaic virus (strain G1)
C;Species: beet necrotic yellow vein mosaic virus
C;Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 08-Apr-1994
C;Accession: D44503
R;Bouzoubaa, S.; Guilley, H.; Jonard, G.; Richards, K.; Putz, C.
J. Gen. Virol. 66, 1553-1564, 1985
A;Title: Nucleotide sequence analysis of RNA-3 and RNA-4 of beet necrotic yellow vein vi
A;Reference number: A44503
A;Accession: D44503
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-135 <BOU>
A;Cross-references: GB:M36894
C;Genetics:
A;Map position: segment 4

  Query Match             42.1%;  Score 48;  DB 2;  Length 135;
  Best Local Similarity   57.1%;  Pred. No. 3.6;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNNDETCE 15
              || |  | :|| ||
Db        115 CYSGVDVLSDELCE 128


RESULT 13
JC5677
RNA4 protein - Beet necrotic yellow vein mosaic virus
C;Species: Beet necrotic yellow vein mosaic virus
C;Date: 11-Nov-1997 #sequence_revision 11-Nov-1997 #text_change 23-Jan-1998
C;Accession: JC5677
R;Yu, J.L.; Han, C.G.; Yan, L.L.; Li, D.W.; Liu, Y.
Acta Microbiol. Sin. 37, 7-14, 1997
A;Title: cDNA cloning, sequence analysis and expression of RNA4 from beet necrotic yello
A;Reference number: JC5677
A;Accession: JC5677
A;Molecule type: mRNA
A;Residues: 1-282 <YUA>

  Query Match             42.1%;  Score 48;  DB 2;  Length 282;
  Best Local Similarity   57.1%;  Pred. No. 7.3;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNNDETCE 15
              || |  | :|| ||
Db        115 CYSGVDVLSDELCE 128


RESULT 14
C44503
p31 protein - beet necrotic yellow vein mosaic virus (strain F2)
C;Species: beet necrotic yellow vein mosaic virus
C;Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 08-Apr-1994
C;Accession: C44503
R;Bouzoubaa, S.; Guilley, H.; Jonard, G.; Richards, K.; Putz, C.
J. Gen. Virol. 66, 1553-1564, 1985
A;Title: Nucleotide sequence analysis of RNA-3 and RNA-4 of beet necrotic yellow vein vi
A;Reference number: A44503
A;Accession: C44503
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-282 <BOU>
A;Cross-references: GB:M36894
C;Genetics:
A;Map position: segment 4

  Query Match             42.1%;  Score 48;  DB 2;  Length 282;
  Best Local Similarity   57.1%;  Pred. No. 7.3;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNNDETCE 15
              || |  | :|| ||
Db        115 CYSGVDVLSDELCE 128


RESULT 15
JQ1965
hypothetical 94K protein - rice stripe virus
C;Species: rice stripe virus
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Jun-2000
```

C;Accession: JQ1965
R;Takahashi, M.; Toriyama, S.; Hamamatsu, C.; Ishihama, A.
J. Gen. Virol. 74, 769-773, 1993
A;Title: Nucleotide sequence and possible ambisense coding strategy of rice stripe virus
A;Reference number: JQ1964; MUID:93224901; PMID:8468559
A;Accession: JQ1965
A;*Molecule type: genomic RNA*
A;Residues: 1-834 <TAK>
A;Cross-references: DDBJ:D13176; NID:g536885; PIDN:BAA02470.1; PID:g1199491
C;Superfamily: rice stripe virus hypothetical 94K protein

Query Match 41.2%; Score 47; DB 2; Length 834;
Best Local Similarity 69.2%; Pred. No. 29;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy          2 CYDGASVNNDETC 14
              ||: ||||  |||
Db         69 CYNRASVNYFETC 81
```

Search completed: December 9, 2003, 14:08:25
Job time : 23 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 14:13:25 ; Search time 23 Seconds
(without alignments)
36.792 Million cell updates/sec

Title: US-09-651-685A-5
Perfect score: 114
Sequence: 1 CCYDGASVNNDETCEQRAAR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 152073

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	35	30.7	16	1	US-08-137-800-14	Sequence 14, Appl
2	35	30.7	16	1	US-08-477-383-14	Sequence 14, Appl
3	35	30.7	16	1	US-08-487-174-14	Sequence 14, Appl
4	35	30.7	16	1	US-08-480-750-14	Sequence 14, Appl
5	35	30.7	19	4	US-09-441-502B-31	Sequence 31, Appl
6	34.5	30.3	18	1	US-08-137-800-32	Sequence 32, Appl
7	34.5	30.3	18	1	US-08-477-383-32	Sequence 32, Appl
8	34.5	30.3	18	1	US-08-487-174-32	Sequence 32, Appl
9	34.5	30.3	18	1	US-08-480-750-32	Sequence 32, Appl
10	34	29.8	15	3	US-08-975-040-3	Sequence 3, Appli
11	34	29.8	19	1	US-08-392-646-18	Sequence 18, Appl
12	34	29.8	19	3	US-08-975-040-2	Sequence 2, Appli
13	34	29.8	19	3	US-08-975-040-19	Sequence 19, Appl
14	34	29.8	19	4	US-09-101-927-17	Sequence 17, Appl
15	34	29.8	19	4	US-09-101-927-19	Sequence 19, Appl
16	34	29.8	19	4	US-09-441-502B-30	Sequence 30, Appl
17	32	28.1	9	4	US-09-106-872A-8	Sequence 8, Appli
18	32	28.1	18	1	US-08-159-340A-32	Sequence 32, Appl
19	31	27.2	16	1	US-08-574-763-3	Sequence 3, Appli
20	31	27.2	20	5	PCT-US95-06726-19	Sequence 19, Appl
21	30	26.3	11	3	US-08-838-413A-14	Sequence 14, Appl
22	29.5	25.9	18	1	US-08-137-800-13	Sequence 13, Appl
23	29.5	25.9	18	1	US-08-137-800-31	Sequence 31, Appl
24	29.5	25.9	18	1	US-08-477-383-13	Sequence 13, Appl
25	29.5	25.9	18	1	US-08-477-383-31	Sequence 31, Appl
26	29.5	25.9	18	1	US-08-487-174-13	Sequence 13, Appl
27	29.5	25.9	18	1	US-08-487-174-31	Sequence 31, Appl
28	29.5	25.9	18	1	US-08-480-750-13	Sequence 13, Appl
29	29.5	25.9	18	1	US-08-480-750-31	Sequence 31, Appl
30	29	25.4	10	1	US-08-421-702A-79	Sequence 79, Appl
31	29	25.4	10	1	US-08-303-052A-32	Sequence 32, Appl
32	29	25.4	10	1	US-08-421-696A-79	Sequence 79, Appl
33	29	25.4	10	1	US-08-421-697A-79	Sequence 79, Appl
34	29	25.4	10	1	US-08-421-698A-79	Sequence 79, Appl
35	29	25.4	10	2	US-08-421-695A-32	Sequence 32, Appl
36	29	25.4	10	3	US-08-461-384B-9	Sequence 9, Appli
37	29	25.4	10	3	US-08-407-207A-6	Sequence 6, Appli
38	29	25.4	13	4	US-09-136-769A-7	Sequence 7, Appli
39	29	25.4	13	4	US-09-136-769A-18	Sequence 18, Appl
40	29	25.4	13	4	US-09-136-769A-26	Sequence 26, Appl
41	29	25.4	19	1	US-08-335-303-20	Sequence 20, Appl
42	29	25.4	19	1	US-08-290-448A-49	Sequence 49, Appl
43	29	25.4	19	1	US-08-290-448A-49	Sequence 49, Appl
44	29	25.4	19	1	US-08-175-069A-49	Sequence 49, Appl
45	29	25.4	19	4	US-08-461-939B-49	Sequence 49, Appl

ALIGNMENTS

```
RESULT 1
US-08-137-800-14
; Sequence 14, Application US/08137800
; Patent No. 5514774
;  GENERAL INFORMATION:
;    APPLICANT:  Olivera, Baldomero M.
;    APPLICANT:  Cruz, Lourdes J.
;    APPLICANT:  Hillyard, David R.
;    APPLICANT:  McIntosh, J. Michael
;    APPLICANT:  Santos, Ameurfina D.
;    TITLE OF INVENTION:  Conotoxin Peptides
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Venable, Baetjer, Howard & Civiletti
;      STREET:  1201 New York Avenue N.W., Suite 1000
;      CITY:  Washington
;      STATE:  DC
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/137,800
;      FILING DATE:  19-OCT-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ihnen, Jeffrey L.
;      REGISTRATION NUMBER:  28,957
;      REFERENCE/DOCKET NUMBER:  24260-104763
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-962-4810
;      TELEFAX:  202-962-8300
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Conus bandanus
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  6..14
;      OTHER INFORMATION:  /note= "Xaa(6) is Pro or
;      OTHER INFORMATION:  Hydroxy-Pro; Xaa(13) is Pro or Hydroxy-Pro;
;      OTHER INFORMATION:  Xaa(14) is Asp or beta-carboxyaspartate"
```

US-08-137-800-14

Query Match 30.7%; Score 35; DB 1; Length 16;
Best Local Similarity 42.9%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETC 14
              | :   ||||   |
Db          3 CSHXACSVNNXXIC 16
```

RESULT 2
US-08-477-383-14
; Sequence 14, Application US/08477383
; Patent No. 5589340
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: Macintosh, J. Michael
; APPLICANT: Santos, Ameurfino S.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,383
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/137,800
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/084,848
; FILING DATE: 29-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-107673
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus bandanus
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 6..14
; OTHER INFORMATION: /note= "Xaa(6) is Pro or
; OTHER INFORMATION: Hydroxy-Pro; Xaa(13) is Pro or Hydroxy-Pro; Xaa(14) is Asp or
; OTHER INFORMATION: beta-carboxyaspartate."
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 16
; OTHER INFORMATION: /note= "The C-terminus is
; OTHER INFORMATION: preferably amidated."

US-08-477-383-14

Query Match 30.7%; Score 35; DB 1; Length 16;
Best Local Similarity 42.9%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETC 14
              | :   ||||   |
Db          3 CSHXACSVNNXXIC 16
```

RESULT 3
US-08-487-174-14
; Sequence 14, Application US/08487174
; Patent No. 5595972
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: Macintosh, J. Michael
; APPLICANT: Santos, Ameurfino S.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,174
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/137,800
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/084,848
; FILING DATE: 29-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-107673
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus bandanus
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 6..14
; OTHER INFORMATION: /note= "Xaa(6) is Pro or
; OTHER INFORMATION: Hydroxy-Pro; Xaa(13) is Pro or Hydroxy-Pro; Xaa(14) is Asp or
; OTHER INFORMATION: beta-carboxyaspartate."
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 16
; OTHER INFORMATION: /note= "The C-terminus is
; OTHER INFORMATION: preferably amidated."

US-08-487-174-14

Query Match 30.7%; Score 35; DB 1; Length 16;
Best Local Similarity 42.9%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETC 14
              | :   ||||   |
Db          3 CSHXACSVNNXXIC 16
```

RESULT 4
US-08-480-750-14
; Sequence 14, Application US/08480750
; Patent No. 5633347
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: Macintosh, J. Michael
; APPLICANT: Santos, Ameurfino S.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,750
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/137,800
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/084,848
; FILING DATE: 29-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-107673
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus bandanus
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 6..14
; OTHER INFORMATION: /note= "Xaa(6) is Pro or
; OTHER INFORMATION: Hydroxy-Pro; Xaa(13) is Pro or Hydroxy-Pro; Xaa(14) is Asp or
; OTHER INFORMATION: beta-carboxyaspartate."
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 16
; OTHER INFORMATION: /note= "The C-terminus is
; OTHER INFORMATION: preferably amidated."

US-08-480-750-14

Query Match 30.7%; Score 35; DB 1; Length 16;
Best Local Similarity 42.9%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETC 14
              | :   ||||   |
Db          3 CSHXACSVNNXXIC 16
```

RESULT 5
US-09-441-502B-31
; Sequence 31, Application US/09441502B
; Patent No. 6455041
; GENERAL INFORMATION:
; APPLICANT: Dunbar, Bonita S.
; TITLE OF INVENTION: IMMUNOGENIC EPITOPES OF THE HUMAN ZONA PELLUCIDA PROTEIN
; FILE REFERENCE: 12231.2USU1
; CURRENT APPLICATION NUMBER: US/09/441,502B
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-502B-31

Query Match 30.7%; Score 35; DB 4; Length 19;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNN 10
              |||    ||:
Db          5 CCYSSEEVNS 14
```

RESULT 6
US-08-137-800-32
; Sequence 32, Application US/08137800
; Patent No. 5514774
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Santos, Ameurfina D.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue N.W., Suite 1000
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,800
; FILING DATE: 19-OCT-1993
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-104763
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 32:

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Conus ermineus
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  3..4
;      OTHER INFORMATION:  /note= "Xaa is Pro or Hydroxy-Pro"
US-08-137-800-32

  Query Match             30.3%;  Score 34.5;  DB 1;  Length 18;
  Best Local Similarity   46.7%;  Pred. No. 45;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 CCYDGA-SVNNDETC 14
              || : | :||| : |
Db          4 CCSNPACNVNNPQIC 18


RESULT 7
US-08-477-383-32
; Sequence 32, Application US/08477383
; Patent No. 5589340
;  GENERAL INFORMATION:
;    APPLICANT:  Olivera, Baldomero M.
;    APPLICANT:  Cruz, Lourdes J.
;    APPLICANT:  Hillyard, David R.
;    APPLICANT:  Macintosh, J. Michael
;    APPLICANT:  Santos, Ameurfino S.
;    TITLE OF INVENTION:  Conotoxin Peptides
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Venable, Baetjer, Howard & Civiletti
;      STREET:  1201 New York Avenue, N.W., Suite 1000
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.A.
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,383
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/137,800
;      FILING DATE:  19-OCT-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/084,848
;      FILING DATE:  29-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ihnen, Jeffrey L.
;      REGISTRATION NUMBER:  28,957
;      REFERENCE/DOCKET NUMBER:  24260-107673
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-962-4810
;      TELEFAX:  202-962-8300
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Conus ermineus
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  3
;      OTHER INFORMATION:  /note= "Xaa is Pro or Hydroxy-Pro."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  18
;      OTHER INFORMATION:  /note= "The C-terminus is
;      OTHER INFORMATION:  preferably amidated."
US-08-477-383-32

  Query Match             30.3%;  Score 34.5;  DB 1;  Length 18;
  Best Local Similarity   46.7%;  Pred. No. 45;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 CCYDGA-SVNNDETC 14
              || : | :||| : |
Db          4 CCSNPACNVNNPQIC 18


RESULT 8
US-08-487-174-32
; Sequence 32, Application US/08487174
; Patent No. 5595972
;  GENERAL INFORMATION:
;    APPLICANT:  Olivera, Baldomero M.
;    APPLICANT:  Cruz, Lourdes J.
;    APPLICANT:  Hillyard, David R.
;    APPLICANT:  Macintosh, J. Michael
;    APPLICANT:  Santos, Ameurfino S.
;    TITLE OF INVENTION:  Conotoxin Peptides
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Venable, Baetjer, Howard & Civiletti
;      STREET:  1201 New York Avenue, N.W., Suite 1000
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.A.
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,174
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/137,800
;      FILING DATE:  19-OCT-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/084,848
;      FILING DATE:  29-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ihnen, Jeffrey L.
;      REGISTRATION NUMBER:  28,957
;      REFERENCE/DOCKET NUMBER:  24260-107673
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-962-4810
;      TELEFAX:  202-962-8300
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Conus ermineus
```

```
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  3
;      OTHER INFORMATION:  /note= "Xaa is Pro or Hydroxy-Pro."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  18
;      OTHER INFORMATION:  /note= "The C-terminus is
;      OTHER INFORMATION:  preferably amidated."
US-08-487-174-32

  Query Match             30.3%;  Score 34.5;  DB 1;  Length 18;
  Best Local Similarity   46.7%;  Pred. No. 45;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 CCYDGA-SVNNDETC 14
              || : | :||| : |
Db          4 CCSNPACNVNNPQIC 18


RESULT 9
US-08-480-750-32
; Sequence 32, Application US/08480750
; Patent No. 5633347
;  GENERAL INFORMATION:
;    APPLICANT:  Olivera, Baldomero M.
;    APPLICANT:  Cruz, Lourdes J.
;    APPLICANT:  Hillyard, David R.
;    APPLICANT:  Macintosh, J. Michael
;    APPLICANT:  Santos, Ameurfino S.
;    TITLE OF INVENTION:  Conotoxin Peptides
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Venable, Baetjer, Howard & Civiletti
;      STREET:  1201 New York Avenue, N.W., Suite 1000
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.A.
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,750
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/137,800
;      FILING DATE:  19-OCT-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/084,848
;      FILING DATE:  29-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ihnen, Jeffrey L.
;      REGISTRATION NUMBER:  28,957
;      REFERENCE/DOCKET NUMBER:  24260-107673
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-962-4810
;      TELEFAX:  202-962-8300
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Conus ermineus
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  3
;      OTHER INFORMATION:  /note= "Xaa is Pro or Hydroxy-Pro."
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  18
;      OTHER INFORMATION:  /note= "The C-terminus is
;      OTHER INFORMATION:  preferably amidated."
US-08-480-750-32

  Query Match             30.3%;  Score 34.5;  DB 1;  Length 18;
  Best Local Similarity   46.7%;  Pred. No. 45;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 CCYDGA-SVNNDETC 14
              || : | :||| : |
Db          4 CCSNPACNVNNPQIC 18


RESULT 10
US-08-975-040-3
; Sequence 3, Application US/08975040
; Patent No. 6251620
;  GENERAL INFORMATION:
;    APPLICANT:  HATADA, MARCOS
;    APPLICANT:  LU, XIAODE
;    APPLICANT:  LAIRD, ELLEN
;    APPLICANT:  KARAS, JENNIFER
;    APPLICANT:  ZOLLER, MARK
;    APPLICANT:  HOLT, DENNIS
;    TITLE OF INVENTION:  MACHINE READABLE STORAGE MEDIUM RELATING
;    TITLE OF INVENTION:  TO ZAP-FAMILY PROTEINS
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DAVID L. BERSTEIN, ARIAD PHARMACEUTICALS,
;      ADDRESSEE:  INC.
;      STREET:  26 LANDSDOWNE STREET
;      CITY:  CAMBRIDGE
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02139
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/975,040
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/605,578
;      FILING DATE:  22-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERSTEIN, DAVID L.
;      REGISTRATION NUMBER:  31,235
;      REFERENCE/DOCKET NUMBER:  ARIAD 347F
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-494-0400
;      TELEFAX:  617-494-1828
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /product= "phosphotyrosine"
;      OTHER INFORMATION:  /label= pTyr
;      OTHER INFORMATION:  /note= "phosphorylated tyrosine"
;    FEATURE:
```

```
;       NAME/KEY:  Modified-site
;       LOCATION:  12
;       OTHER INFORMATION:  /product= "phosphotyrosine"
;       OTHER INFORMATION:  /label= pTyr
;       OTHER INFORMATION:  /note= "phosphorylated tyrosine"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  1
;       OTHER INFORMATION:  /product= "Acetylated"
;       OTHER INFORMATION:  /label= Ac
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  15
;       OTHER INFORMATION:  /product= "amidated"
;       OTHER INFORMATION:  /label= NH2
US-08-975-040-3

  Query Match             29.8%;  Score 34;  DB 3;  Length 15;
  Best Local Similarity   53.8%;  Pred. No. 44;
  Matches    7;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 YDGASVNNDETCE 15
              | | |  | || |
Db          1 YTGLSTRNQETYE 13


RESULT 11
US-08-392-646-18
; Sequence 18, Application US/08392646
; Patent No. 5710129
;  GENERAL INFORMATION:
;    APPLICANT:  LYNCH, Berkley A.
;    APPLICANT:  WEIGELE, Manfred
;    TITLE OF INVENTION:  NEW INHIBITORS OF SH2-MEDIATED PROCESSES
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARIAD Pharmaceuticals, Inc.
;      STREET:  26 Landsdowne Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02139-4234
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/392,646
;      FILING DATE:  23-FEB-1995
;      CLASSIFICATION:  560
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERSTEIN, David L.
;      REGISTRATION NUMBER:  31,235
;      REFERENCE/DOCKET NUMBER:  ARIAD 337
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-494-0400
;      TELEFAX:  617-494-0208
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  4
;      OTHER INFORMATION:  /product= "phosphotyrosine"
;      OTHER INFORMATION:  /label= YPO4
;      OTHER INFORMATION:  /note= "phosphorylated tyrosine"
;    FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  15
;       OTHER INFORMATION:  /product= "phosphotyrosine"
;       OTHER INFORMATION:  /label= YPO4
;       OTHER INFORMATION:  /note= "phosphorylated tyrosine"
US-08-392-646-18

  Query Match             29.8%;  Score 34;  DB 1;  Length 19;
  Best Local Similarity   53.8%;  Pred. No. 56;
  Matches    7;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 YDGASVNNDETCE 15
              | | |  | || |
Db          4 YTGLSTRNQETYE 16


RESULT 12
US-08-975-040-2
; Sequence 2, Application US/08975040
; Patent No. 6251620
;  GENERAL INFORMATION:
;    APPLICANT:  HATADA, MARCOS
;    APPLICANT:  LU, XIAODE
;    APPLICANT:  LAIRD, ELLEN
;    APPLICANT:  KARAS, JENNIFER
;    APPLICANT:  ZOLLER, MARK
;    APPLICANT:  HOLT, DENNIS
;    TITLE OF INVENTION:  MACHINE READABLE STORAGE MEDIUM RELATING
;    TITLE OF INVENTION:  TO ZAP-FAMILY PROTEINS
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DAVID L. BERSTEIN, ARIAD PHARMACEUTICALS,
;      ADDRESSEE:  INC.
;      STREET:  26 LANDSDOWNE STREET
;      CITY:  CAMBRIDGE
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02139
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/975,040
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/605,578
;      FILING DATE:  22-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERSTEIN, DAVID L.
;      REGISTRATION NUMBER:  31,235
;      REFERENCE/DOCKET NUMBER:  ARIAD 347F
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-494-0400
;      TELEFAX:  617-494-1828
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  4
;      OTHER INFORMATION:  /product= "phosphotyrosine"
;      OTHER INFORMATION:  /label= pTyr
;      OTHER INFORMATION:  /note= "phosphorylated tyrosine"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  15
```

```
;       OTHER INFORMATION:  /product= "phosphotyrosine"
;       OTHER INFORMATION:  /label= pTyr
;       OTHER INFORMATION:  /note= "phosphorylated tyrosine"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  1
;       OTHER INFORMATION:  /product= "Acetylated"
;       OTHER INFORMATION:  /label= Ac
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  19
;       OTHER INFORMATION:  /product= "amidated"
;       OTHER INFORMATION:  /label= NH2
US-08-975-040-2

  Query Match             29.8%;  Score 34;  DB 3;  Length 19;
  Best Local Similarity   53.8%;  Pred. No. 56;
  Matches    7;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 YDGASVNNDETCE 15
              | | |   | || |
Db          4 YTGLSTRNQETYE 16


RESULT 13
US-08-975-040-19
; Sequence 19, Application US/08975040
; Patent No. 6251620
;  GENERAL INFORMATION:
;    APPLICANT:  HATADA, MARCOS
;    APPLICANT:  LU, XIAODE
;    APPLICANT:  LAIRD, ELLEN
;    APPLICANT:  KARAS, JENNIFER
;    APPLICANT:  ZOLLER, MARK
;    APPLICANT:  HOLT, DENNIS
;    TITLE OF INVENTION:  MACHINE READABLE STORAGE MEDIUM RELATING
;    TITLE OF INVENTION:  TO ZAP-FAMILY PROTEINS
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DAVID L. BERSTEIN, ARIAD PHARMACEUTICALS,
;      ADDRESSEE:  INC.
;      STREET:  26 LANDSDOWNE STREET
;      CITY:  CAMBRIDGE
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02139
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/975,040
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/605,578
;      FILING DATE:  22-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BERSTEIN, DAVID L.
;      REGISTRATION NUMBER:  31,235
;      REFERENCE/DOCKET NUMBER:  ARIAD 347F
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-494-0400
;      TELEFAX:  617-494-1828
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  4
;      OTHER INFORMATION:  /product= "phosphotyrosine"
;      OTHER INFORMATION:  /label= YPO4
;      OTHER INFORMATION:  /note= "phosphorylated tyrosine"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  15
;      OTHER INFORMATION:  /product= "phosphotyrosine"
;      OTHER INFORMATION:  /label= YPO4
;      OTHER INFORMATION:  /note= "phosphorylated tyrosine"
US-08-975-040-19

  Query Match             29.8%;  Score 34;  DB 3;  Length 19;
  Best Local Similarity   53.8%;  Pred. No. 56;
  Matches    7;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 YDGASVNNDETCE 15
              | | |   | || |
Db          4 YTGLSTRNQETYE 16


RESULT 14
US-09-101-927-17
; Sequence 17, Application US/09101927
; Patent No. 6303319
; GENERAL INFORMATION:
;  APPLICANT: Rickles, Richard J
;  TITLE OF INVENTION: Cell-Based Assay
;  FILE REFERENCE: 336A PCT/US
;  CURRENT APPLICATION NUMBER: US/09/101,927
;  CURRENT FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: PCT/US97/026735
;  EARLIER FILING DATE: 1997-02-21
;  EARLIER APPLICATION NUMBER: 60/012,218
;  EARLIER FILING DATE: 1996-02-23
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (4)
;   OTHER INFORMATION: PHOSPHORYLATION
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (15)
;   OTHER INFORMATION: PHOSPHORYLATION
US-09-101-927-17

  Query Match             29.8%;  Score 34;  DB 4;  Length 19;
  Best Local Similarity   53.8%;  Pred. No. 56;
  Matches    7;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 YDGASVNNDETCE 15
              | | |   | || |
Db          4 YTGLSTRNQETYE 16


RESULT 15
US-09-101-927-19
; Sequence 19, Application US/09101927
; Patent No. 6303319
; GENERAL INFORMATION:
;  APPLICANT: Rickles, Richard J
;  TITLE OF INVENTION: Cell-Based Assay
;  FILE REFERENCE: 336A PCT/US
;  CURRENT APPLICATION NUMBER: US/09/101,927
;  CURRENT FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: PCT/US97/026735
```

; EARLIER FILING DATE: 1997-02-21
; EARLIER APPLICATION NUMBER: 60/012,218
; EARLIER FILING DATE: 1996-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (4)
; OTHER INFORMATION: PHOSPHORYLATION
US-09-101-927-19

Query Match 29.8%; Score 34; DB 4; Length 19;
Best Local Similarity 53.8%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          3 YDGASVNNDETCE 15
              | | |  | || |
Db          4 YTGLSTRNQETYE 16
```

Search completed: December 9, 2003, 14:16:30
Job time : 24 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 14:15:25 ; Search time 30 Seconds
(without alignments)
123.989 Million cell updates/sec

Title: US-09-651-685A-5
Perfect score: 114
Sequence: 1 CCYDGASVNNDETCEQRAAR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 145776

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	114	100.0	20	10	US-09-878-603-5	Sequence 5, Appli
2	111	97.4	20	10	US-09-878-603-68	Sequence 68, Appl
3	110	96.5	20	10	US-09-878-603-67	Sequence 67, Appl
4	110	96.5	20	10	US-09-878-603-69	Sequence 69, Appl
5	110	96.5	20	10	US-09-878-603-70	Sequence 70, Appl
6	110	96.5	20	10	US-09-878-603-71	Sequence 71, Appl
7	110	96.5	20	10	US-09-878-603-73	Sequence 73, Appl
8	109	95.6	19	10	US-09-878-603-54	Sequence 54, Appl
9	108	94.7	20	10	US-09-878-603-64	Sequence 64, Appl
10	108	94.7	20	10	US-09-878-603-65	Sequence 65, Appl
11	106	93.0	20	10	US-09-878-603-72	Sequence 72, Appl
12	105	92.1	18	10	US-09-878-603-55	Sequence 55, Appl
13	105	92.1	19	10	US-09-878-603-59	Sequence 59, Appl
14	102	89.5	20	10	US-09-878-603-66	Sequence 66, Appl
15	101	88.6	17	10	US-09-878-603-56	Sequence 56, Appl
16	98	86.0	20	10	US-09-878-603-74	Sequence 74, Appl
17	96	84.2	16	10	US-09-878-603-57	Sequence 57, Appl
18	96	84.2	18	10	US-09-878-603-60	Sequence 60, Appl
19	91	79.8	15	10	US-09-878-603-58	Sequence 58, Appl
20	89	78.1	17	10	US-09-878-603-61	Sequence 61, Appl
21	83	72.8	16	10	US-09-878-603-62	Sequence 62, Appl
22	77	67.5	15	10	US-09-878-603-63	Sequence 63, Appl
23	51	44.7	20	10	US-09-878-603-2	Sequence 2, Appli
24	36	31.6	15	12	US-10-256-277-46	Sequence 46, Appl
25	36	31.6	15	15	US-10-209-323-46	Sequence 46, Appl
26	35.5	31.1	19	15	US-10-072-602B-407	Sequence 407, App
27	35	30.7	16	15	US-10-072-602B-444	Sequence 444, App
28	34.5	30.3	19	15	US-10-072-602B-604	Sequence 604, App
29	34	29.8	18	15	US-10-142-238A-3	Sequence 3, Appli
30	34	29.8	20	10	US-09-321-801-5	Sequence 5, Appli
31	33	28.9	16	15	US-10-072-602B-617	Sequence 617, App
32	33	28.9	17	15	US-10-072-602B-403	Sequence 403, App
33	33	28.9	17	15	US-10-072-602B-404	Sequence 404, App
34	33	28.9	17	15	US-10-072-602B-602	Sequence 602, App
35	33	28.9	17	15	US-10-072-602B-603	Sequence 603, App
36	32	28.1	6	10	US-09-878-603-15	Sequence 15, Appl
37	31	27.2	17	15	US-10-072-602B-606	Sequence 606, App
38	31	27.2	20	12	US-10-159-339-91	Sequence 91, Appl
39	30.5	26.8	16	12	US-10-193-109-16	Sequence 16, Appl
40	30.5	26.8	16	15	US-10-084-994-16	Sequence 16, Appl
41	30	26.3	12	12	US-10-256-277-44	Sequence 44, Appl
42	30	26.3	12	15	US-10-209-323-44	Sequence 44, Appl
43	30	26.3	18	12	US-10-256-277-48	Sequence 48, Appl
44	30	26.3	18	15	US-10-209-323-48	Sequence 48, Appl
45	29	25.4	9	10	US-09-252-150-78	Sequence 78, Appl

ALIGNMENTS

RESULT 1
US-09-878-603-5
; Sequence 5, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-878-603-5

Query Match 100.0%; Score 114; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||||||||||||||||
Db 1 CCYDGASVNNDETCEQRAAR 20

RESULT 2
US-09-878-603-68
; Sequence 68, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.

; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-68

Query Match 97.4%; Score 111; DB 10; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.9e-10;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||:||||||||||||
Db 1 CCYDGATVNNDETCEQRAAR 20

RESULT 3
US-09-878-603-67
; Sequence 67, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 67
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-67

Query Match 96.5%; Score 110; DB 10; Length 20;
Best Local Similarity 95.0%; Pred. No. 4.2e-10;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||||||||:|||||||
Db 1 CCYDGASVNNDESCEQRAAR 20

RESULT 4
US-09-878-603-69
; Sequence 69, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-69

Query Match 96.5%; Score 110; DB 10; Length 20;
Best Local Similarity 95.0%; Pred. No. 4.2e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||| ||||||||||||||
Db 1 CCYDGVSVNNDETCEQRAAR 20

RESULT 5
US-09-878-603-70
; Sequence 70, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-70

Query Match 96.5%; Score 110; DB 10; Length 20;
Best Local Similarity 95.0%; Pred. No. 4.2e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||| ||||||||||||
Db 1 CCYDGASANNDETCEQRAAR 20

RESULT 6
US-09-878-603-71
; Sequence 71, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74

```
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
US-09-878-603-71

  Query Match             96.5%;  Score 110;  DB 10;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 4.2e-10;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||||||||||||||| ||
Db          1 CCYDGASVNNDETCEQRVAR 20


RESULT 7
US-09-878-603-73
; Sequence 73, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
;  APPLICANT: Ward, Peter A.
;  APPLICANT:  Huber-Lang, Markus
;  APPLICANT:  Sarma, Vidya
;  APPLICANT:  Czermak, Boris
;  TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
;  FILE REFERENCE: UM-03783
;  CURRENT APPLICATION NUMBER: US/09/878,603
;  CURRENT FILING DATE:  2001-06-11
;  PRIOR APPLICATION NUMBER: 09/387,671
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
US-09-878-603-73

  Query Match             96.5%;  Score 110;  DB 10;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 4.2e-10;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||||||||||||||| |
Db          1 CCYDGASVNNDETCEQRAVR 20


RESULT 8
US-09-878-603-54
; Sequence 54, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
;  APPLICANT: Ward, Peter A.
;  APPLICANT:  Huber-Lang, Markus
;  APPLICANT:  Sarma, Vidya
;  APPLICANT:  Czermak, Boris
;  TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
;  FILE REFERENCE: UM-03783
;  CURRENT APPLICATION NUMBER: US/09/878,603
;  CURRENT FILING DATE:  2001-06-11
;  PRIOR APPLICATION NUMBER: 09/387,671
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
US-09-878-603-54

  Query Match             95.6%;  Score 109;  DB 10;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAA 19
              |||||||||||||||||||
Db          1 CCYDGASVNNDETCEQRAA 19


RESULT 9
US-09-878-603-64
; Sequence 64, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
;  APPLICANT: Ward, Peter A.
;  APPLICANT:  Huber-Lang, Markus
;  APPLICANT:  Sarma, Vidya
;  APPLICANT:  Czermak, Boris
;  TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
;  FILE REFERENCE: UM-03783
;  CURRENT APPLICATION NUMBER: US/09/878,603
;  CURRENT FILING DATE:  2001-06-11
;  PRIOR APPLICATION NUMBER: 09/387,671
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
US-09-878-603-64

  Query Match             94.7%;  Score 108;  DB 10;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 8.4e-10;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||| ||||||||||||||||
Db          1 CCYQGASVNNDETCEQRAAR 20


RESULT 10
US-09-878-603-65
; Sequence 65, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
;  APPLICANT: Ward, Peter A.
;  APPLICANT:  Huber-Lang, Markus
;  APPLICANT:  Sarma, Vidya
;  APPLICANT:  Czermak, Boris
;  TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
;  FILE REFERENCE: UM-03783
;  CURRENT APPLICATION NUMBER: US/09/878,603
;  CURRENT FILING DATE:  2001-06-11
;  PRIOR APPLICATION NUMBER: 09/387,671
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 65
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
US-09-878-603-65

  Query Match             94.7%;  Score 108;  DB 10;  Length 20;
```

Best Local Similarity 95.0%; Pred. No. 8.4e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||||||| |||||||||
Db          1 CCYDGASVNNQETCEQRAAR 20
```

RESULT 11
US-09-878-603-72
; Sequence 72, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 72
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-72

Query Match 93.0%; Score 106; DB 10; Length 20;
Best Local Similarity 90.0%; Pred. No. 1.7e-09;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||||||||||||||  |
Db          1 CCYDGASVNNDETCEQRVVR 20
```

RESULT 12
US-09-878-603-55
; Sequence 55, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-55

Query Match 92.1%; Score 105; DB 10; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.1e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRA 18
              ||||||||||||||||||
Db          1 CCYDGASVNNDETCEQRA 18
```

RESULT 13
US-09-878-603-59
; Sequence 59, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-59

Query Match 92.1%; Score 105; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 CYDGASVNNDETCEQRAAR 20
              |||||||||||||||||||
Db          1 CYDGASVNNDETCEQRAAR 19
```

RESULT 14
US-09-878-603-66
; Sequence 66, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-66

Query Match 89.5%; Score 102; DB 10; Length 20;
Best Local Similarity 90.0%; Pred. No. 6.9e-09;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQRAAR 20
              ||| |||||| |||||||||
Db          1 CCYQGASVNNQETCEQRAAR 20
```

RESULT 15
US-09-878-603-56

; Sequence 56, Application US/09878603
; Patent No. US20020165138A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-878-603-56

Query Match 88.6%; Score 101; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 8.2e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCYDGASVNNDETCEQR 17
              |||||||||||||||||
Db          1 CCYDGASVNNDETCEQR 17
```

Search completed: December 9, 2003, 14:20:39
Job time : 30 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 14:08:29 ; Search time 41 Seconds
(without alignments)
77.428 Million cell updates/sec

Title: US-09-651-685A-5
Perfect score: 114
Sequence: 1 CCYDGASVNNDETCEQRAAR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 400068

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	114	100.0	20	22	AAB74055	Human C5a peptide
2	111	97.4	20	22	AAB74111	C-terminal truncat
3	110	96.5	20	22	AAB74110	C-terminal truncat
4	110	96.5	20	22	AAB74112	C-terminal truncat
5	110	96.5	20	22	AAB74113	C-terminal truncat
6	110	96.5	20	22	AAB74114	C-terminal truncat
7	110	96.5	20	22	AAB74116	C-terminal truncat
8	109	95.6	19	22	AAB74097	C-terminal truncat
9	109	95.6	20	22	AAB74120	Human C5a peptide
10	108	94.7	20	22	AAB74107	C-terminal truncat
11	108	94.7	20	22	AAB74108	C-terminal truncat
12	106	93.0	20	22	AAB74115	C-terminal truncat
13	105	92.1	18	22	AAB74098	C-terminal truncat
14	105	92.1	19	22	AAB74102	C-terminal truncat
15	102	89.5	20	22	AAB74109	C-terminal truncat
16	101	88.6	17	22	AAB74099	C-terminal truncat
17	98	86.0	20	22	AAB74117	C-terminal truncat
18	96	84.2	16	22	AAB74100	C-terminal truncat
19	96	84.2	18	22	AAB74103	C-terminal truncat
20	91	79.8	15	22	AAB74101	C-terminal truncat
21	89	78.1	17	22	AAB74104	C-terminal truncat
22	83	72.8	16	22	AAB74105	C-terminal truncat
23	77	67.5	15	22	AAB74106	C-terminal truncat
24	65	57.0	12	24	ABP96210	Binding epitope of
25	51	44.7	20	22	AAB74052	Rat C5a peptide fr
26	47	41.2	12	24	ABP96211	Binding epitope of
27	40.5	35.5	18	21	AAB21536	Cone snail alpha-c
28	38.5	33.8	16	21	AAB21533	Cone snail alpha-c
29	38.5	33.8	20	21	AAB21498	Cone snail alpha-c
30	38	33.3	16	16	AAR75274	A-lineage conotoxi
31	38	33.3	16	21	AAB21522	Cone snail alpha-c
32	36	31.6	15	23	AAU80449	Positive cell sele
33	35.5	31.1	17	21	AAB21549	Cone snail alpha-c
34	35.5	31.1	19	23	ABG99638	Conus sp conotoxin
35	35	30.7	16	18	AAW24881	Predatory cone sna
36	35	30.7	16	18	AAW12736	A-lineage conotoxi
37	35	30.7	16	23	ABG99663	Conus sp conotoxin
38	35	30.7	19	24	ABU13434	Zona pellucida pro
39	34.5	30.3	18	16	AAR75292	A-lineage conotoxi
40	34.5	30.3	18	18	AAW24897	Predatory cone sna
41	34.5	30.3	18	18	AAW12751	A-lineage conotoxi
42	34.5	30.3	19	23	ABG99819	Conus sp conotoxin
43	34	29.8	15	18	AAW30137	Tandem SH2 ligand
44	34	29.8	19	15	AAR56639	Tyrosine activatio
45	34	29.8	19	15	AAR56633	Tyrosine activatio

ALIGNMENTS

RESULT 1
AAB74055
ID AAB74055 standard; Peptide; 20 AA.
XX
AC AAB74055;
XX
DT 16-MAY-2001 (first entry)
XX
DE Human C5a peptide fragment #2.
XX
KW Human; C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Homo sapiens.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
DR N-PSDB; AAF75793.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and

PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Claim 8; Page 26; 84pp; English.
XX
CC The present sequence is a peptide fragment of human complement component
CC C5a (the full-length sequence is given in AAB74053). The present
CC invention relates to an antibody specific for the present sequence. The
CC C5a-antibody can be used in a therapeutic composition, which is useful
CC for treating a subject suffering from bacterial infection, e.g.
CC sinusitis, meningitis, respiratory, gastrointestinal or urinary tract
CC infections or infections in wounds. In addition, the C5a antibody can
CC be used for treating sepsis. C5a is also known as anaphylatoxin.
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 114; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.4e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||||||||||||||||
Db 1 CCYDGASVNNDETCEQRAAR 20

RESULT 2
AAB74111
ID AAB74111 standard; Peptide; 20 AA.
XX
AC AAB74111;
XX
DT 16-MAY-2001 (first entry)
XX
DE C-terminal truncated C5a peptide #49.
XX
KW C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Unidentified.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Disclosure; Page 30; 84pp; English.
XX
CC The present sequence is a C-terminal truncated C5a peptide fragment. The
CC present invention relates to an antibody specific for the present
CC sequence. The C5a-antibody can be used in a therapeutic composition,
CC which is useful for treating a subject suffering from bacterial
CC infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC urinary tract infections or infections in wounds. In addition, the C5a
CC antibody can be used for treating sepsis. C5a is also known as
CC anaphylatoxin.
XX
SQ Sequence 20 AA;

Query Match 97.4%; Score 111; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.6e-10;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||:|||||||||||||
Db 1 CCYDGATVNNDETCEQRAAR 20

RESULT 3
AAB74110
ID AAB74110 standard; Peptide; 20 AA.
XX
AC AAB74110;
XX
DT 16-MAY-2001 (first entry)
XX
DE C-terminal truncated C5a peptide #48.
XX
KW C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Unidentified.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Disclosure; Page 30; 84pp; English.
XX
CC The present sequence is a C-terminal truncated C5a peptide fragment. The
CC present invention relates to an antibody specific for the present
CC sequence. The C5a-antibody can be used in a therapeutic composition,
CC which is useful for treating a subject suffering from bacterial
CC infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC urinary tract infections or infections in wounds. In addition, the C5a
CC antibody can be used for treating sepsis. C5a is also known as
CC anaphylatoxin.
XX
SQ Sequence 20 AA;

Query Match 96.5%; Score 110; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.3e-10;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||||||||:|||||||
Db 1 CCYDGASVNNDESCEQRAAR 20

RESULT 4
AAB74112
ID AAB74112 standard; Peptide; 20 AA.
XX
AC AAB74112;
XX
DT 16-MAY-2001 (first entry)
XX
DE C-terminal truncated C5a peptide #50.

XX
KW C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Unidentified.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Disclosure; Page 30; 84pp; English.
XX
CC The present sequence is a C-terminal truncated C5a peptide fragment. The
CC present invention relates to an antibody specific for the present
CC sequence. The C5a-antibody can be used in a therapeutic composition,
CC which is useful for treating a subject suffering from bacterial
CC infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC urinary tract infections or infections in wounds. In addition, the C5a
CC antibody can be used for treating sepsis. C5a is also known as
CC anaphylatoxin.
XX
SQ Sequence 20 AA;

Query Match 96.5%; Score 110; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.3e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||| ||||||||||||||
Db 1 CCYDGVSVNNDETCEQRAAR 20

RESULT 5
AAB74113
ID AAB74113 standard; Peptide; 20 AA.
XX
AC AAB74113;
XX
DT 16-MAY-2001 (first entry)
XX
DE C-terminal truncated C5a peptide #51.
XX
KW C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Unidentified.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Disclosure; Page 30; 84pp; English.
XX
CC The present sequence is a C-terminal truncated C5a peptide fragment. The
CC present invention relates to an antibody specific for the present
CC sequence. The C5a-antibody can be used in a therapeutic composition,
CC which is useful for treating a subject suffering from bacterial
CC infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC urinary tract infections or infections in wounds. In addition, the C5a
CC antibody can be used for treating sepsis. C5a is also known as
SQ Sequence 20 AA;

Query Match 96.5%; Score 110; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.3e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||| ||||||||||||
Db 1 CCYDGASANNDETCEQRAAR 20

RESULT 6
AAB74114
ID AAB74114 standard; Peptide; 20 AA.
XX
AC AAB74114;
XX
DT 16-MAY-2001 (first entry)
XX
DE C-terminal truncated C5a peptide #52.
XX
KW C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Unidentified.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Disclosure; Page 30; 84pp; English.
XX
CC The present sequence is a C-terminal truncated C5a peptide fragment. The
CC present invention relates to an antibody specific for the present
CC sequence. The C5a-antibody can be used in a therapeutic composition,
CC which is useful for treating a subject suffering from bacterial
CC infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC urinary tract infections or infections in wounds. In addition, the C5a
CC antibody can be used for treating sepsis. C5a is also known as

```
CC   anaphylatoxin.
XX
SQ   Sequence   20 AA;

  Query Match             96.5%;  Score 110;  DB 22;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 2.3e-10;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||||||||||||||| ||
Db          1 CCYDGASVNNDETCEQRVAR 20


RESULT 7
AAB74116
ID   AAB74116 standard; Peptide; 20 AA.
XX
AC   AAB74116;
XX
DT   16-MAY-2001  (first entry)
XX
DE   C-terminal truncated C5a peptide #54.
XX
KW   C5a; complement; antibody; bacterial infection; sinusitis;
KW   meningitis; respiratory; gastrointestinal; urinary tract infection;
KW   wound; anaphylatoxin; sepsis.
XX
OS   Unidentified.
XX
PN   WO200115731-A1.
XX
PD   08-MAR-2001.
XX
PF   31-AUG-2000; 2000WO-US24219.
XX
PR   31-AUG-1999;   99US-0387671.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Ward PA,  Huber-Lang M,  Sarma V;
XX
DR   WPI; 2001-226665/23.
XX
PT   Compositions for treating blood-borne and toxin mediated diseases and
PT   treatment of sepsis in humans and other animals comprises anti-C5a
PT   antibodies generated against C-terminal truncated C5a peptides  -
XX
PS   Disclosure; Page 30; 84pp; English.
XX
CC   The present sequence is a C-terminal truncated C5a peptide fragment. The
CC   present invention relates to an antibody specific for the present
CC   sequence. The C5a-antibody can be used in a therapeutic composition,
CC   which is useful for treating a subject suffering from bacterial
CC   infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC   urinary tract infections or infections in wounds. In addition, the C5a
CC   antibody can be used for treating sepsis. C5a is also known as
CC   anaphylatoxin.
XX
SQ   Sequence   20 AA;

  Query Match             96.5%;  Score 110;  DB 22;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 2.3e-10;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAAR 20
              |||||||||||||||||| |
Db          1 CCYDGASVNNDETCEQRAVR 20


RESULT 8
AAB74097
ID   AAB74097 standard; Peptide; 19 AA.
XX
AC   AAB74097;
XX
DT   16-MAY-2001  (first entry)
XX
DE   C-terminal truncated C5a peptide #35.
XX
KW   C5a; complement; antibody; bacterial infection; sinusitis;
KW   meningitis; respiratory; gastrointestinal; urinary tract infection;
KW   wound; anaphylatoxin; sepsis.
XX
OS   Unidentified.
XX
PN   WO200115731-A1.
XX
PD   08-MAR-2001.
XX
PF   31-AUG-2000; 2000WO-US24219.
XX
PR   31-AUG-1999;   99US-0387671.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Ward PA,  Huber-Lang M,  Sarma V;
XX
DR   WPI; 2001-226665/23.
XX
PT   Compositions for treating blood-borne and toxin mediated diseases and
PT   treatment of sepsis in humans and other animals comprises anti-C5a
PT   antibodies generated against C-terminal truncated C5a peptides  -
XX
PS   Disclosure; Page 30; 84pp; English.
XX
CC   The present sequence is a C-terminal truncated C5a peptide fragment. The
CC   present invention relates to an antibody specific for the present
CC   sequence. The C5a-antibody can be used in a therapeutic composition,
CC   which is useful for treating a subject suffering from bacterial
CC   infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC   urinary tract infections or infections in wounds. In addition, the C5a
CC   antibody can be used for treating sepsis. C5a is also known as
CC   anaphylatoxin.
XX
SQ   Sequence   19 AA;

  Query Match             95.6%;  Score 109;  DB 22;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETCEQRAA 19
              |||||||||||||||||||
Db          1 CCYDGASVNNDETCEQRAA 19


RESULT 9
AAB74120
ID   AAB74120 standard; Peptide; 20 AA.
XX
AC   AAB74120;
XX
DT   16-MAY-2001  (first entry)
XX
DE   Human C5a peptide fragment #7.
XX
KW   Human; C5a; complement; antibody; bacterial infection; sinusitis;
KW   meningitis; respiratory; gastrointestinal; urinary tract infection;
KW   wound; anaphylatoxin; sepsis.
XX
OS   Homo sapiens.
XX
PN   WO200115731-A1.
XX
PD   08-MAR-2001.
XX
PF   31-AUG-2000; 2000WO-US24219.
```

XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Disclosure; Fig 15; 84pp; English.
XX
CC The present sequence is a peptide fragment of human complement component
CC C5a (the full-length sequence is given in AAB74119). The present
CC invention relates to an antibody specific for the present sequence. The
CC C5a-antibody can be used in a therapeutic composition, which is useful
CC for treating a subject suffering from bacterial infection, e.g.
CC sinusitis, meningitis, respiratory, gastrointestinal or urinary tract
CC infections or infections in wounds. In addition, the C5a antibody can
CC be used for treating sepsis. C5a is also known as anaphylatoxin.
XX
SQ Sequence 20 AA;

Query Match 95.6%; Score 109; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 3.3e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||| ||||||||||||
Db 1 CCYDGACVNNDETCEQRAAR 20

RESULT 10
AAB74107
ID AAB74107 standard; Peptide; 20 AA.
XX
AC AAB74107;
XX
DT 16-MAY-2001 (first entry)
XX
DE C-terminal truncated C5a peptide #45.
XX
KW C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Unidentified.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Disclosure; Page 30; 84pp; English.
XX
CC The present sequence is a C-terminal truncated C5a peptide fragment. The
CC present invention relates to an antibody specific for the present
CC sequence. The C5a-antibody can be used in a therapeutic composition,
CC which is useful for treating a subject suffering from bacterial
CC infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC urinary tract infections or infections in wounds. In addition, the C5a
CC antibody can be used for treating sepsis. C5a is also known as
CC anaphylatoxin.
XX
SQ Sequence 20 AA;

Query Match 94.7%; Score 108; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 4.7e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||| |||||||||||||||
Db 1 CCYQGASVNNDETCEQRAAR 20

RESULT 11
AAB74108
ID AAB74108 standard; Peptide; 20 AA.
XX
AC AAB74108;
XX
DT 16-MAY-2001 (first entry)
XX
DE C-terminal truncated C5a peptide #46.
XX
KW C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Unidentified.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Disclosure; Page 30; 84pp; English.
XX
CC The present sequence is a C-terminal truncated C5a peptide fragment. The
CC present invention relates to an antibody specific for the present
CC sequence. The C5a-antibody can be used in a therapeutic composition,
CC which is useful for treating a subject suffering from bacterial
CC infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC urinary tract infections or infections in wounds. In addition, the C5a
CC antibody can be used for treating sepsis. C5a is also known as
CC anaphylatoxin.
XX
SQ Sequence 20 AA;

Query Match 94.7%; Score 108; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 4.7e-10;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
|||||||||| |||||||||
Db 1 CCYDGASVNNQETCEQRAAR 20

RESULT 12
AAB74115
ID AAB74115 standard; Peptide; 20 AA.
XX
AC AAB74115;
XX
DT 16-MAY-2001 (first entry)
XX
DE C-terminal truncated C5a peptide #53.
XX
KW C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Unidentified.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Disclosure; Page 30; 84pp; English.
XX
CC The present sequence is a C-terminal truncated C5a peptide fragment. The
CC present invention relates to an antibody specific for the present
CC sequence. The C5a-antibody can be used in a therapeutic composition,
CC which is useful for treating a subject suffering from bacterial
CC infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC urinary tract infections or infections in wounds. In addition, the C5a
CC antibody can be used for treating sepsis. C5a is also known as
CC anaphylatoxin.
XX
SQ Sequence 20 AA;

Query Match 93.0%; Score 106; DB 22; Length 20;
Best Local Similarity 90.0%; Pred. No. 9.7e-10;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||||||||||||||||| |
Db 1 CCYDGASVNNDETCEQRVVR 20

RESULT 13
AAB74098
ID AAB74098 standard; Peptide; 18 AA.
XX
AC AAB74098;
XX
DT 16-MAY-2001 (first entry)
XX
DE C-terminal truncated C5a peptide #36.
XX
KW C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Unidentified.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Disclosure; Page 30; 84pp; English.
XX
CC The present sequence is a C-terminal truncated C5a peptide fragment. The
CC present invention relates to an antibody specific for the present
CC sequence. The C5a-antibody can be used in a therapeutic composition,
CC which is useful for treating a subject suffering from bacterial
CC infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC urinary tract infections or infections in wounds. In addition, the C5a
CC antibody can be used for treating sepsis. C5a is also known as
CC anaphylatoxin.
XX
SQ Sequence 18 AA;

Query Match 92.1%; Score 105; DB 22; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRA 18
||||||||||||||||||
Db 1 CCYDGASVNNDETCEQRA 18

RESULT 14
AAB74102
ID AAB74102 standard; Peptide; 19 AA.
XX
AC AAB74102;
XX
DT 16-MAY-2001 (first entry)
XX
DE C-terminal truncated C5a peptide #40.
XX
KW C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Unidentified.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -

XX
PS Disclosure; Page 30; 84pp; English.
XX
CC The present sequence is a C-terminal truncated C5a peptide fragment. The
CC present invention relates to an antibody specific for the present
CC sequence. The C5a-antibody can be used in a therapeutic composition,
CC which is useful for treating a subject suffering from bacterial
CC infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC urinary tract infections or infections in wounds. In addition, the C5a
CC antibody can be used for treating sepsis. C5a is also known as
CC anaphylatoxin.
XX
SQ Sequence 19 AA;

Query Match 92.1%; Score 105; DB 22; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CYDGASVNNDETCEQRAAR 20
||||||||||||||||||||
Db 1 CYDGASVNNDETCEQRAAR 19

RESULT 15
AAB74109
ID AAB74109 standard; Peptide; 20 AA.
XX
AC AAB74109;
XX
DT 16-MAY-2001 (first entry)
XX
DE C-terminal truncated C5a peptide #47.
XX
KW C5a; complement; antibody; bacterial infection; sinusitis;
KW meningitis; respiratory; gastrointestinal; urinary tract infection;
KW wound; anaphylatoxin; sepsis.
XX
OS Unidentified.
XX
PN WO200115731-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US24219.
XX
PR 31-AUG-1999; 99US-0387671.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Ward PA, Huber-Lang M, Sarma V;
XX
DR WPI; 2001-226665/23.
XX
PT Compositions for treating blood-borne and toxin mediated diseases and
PT treatment of sepsis in humans and other animals comprises anti-C5a
PT antibodies generated against C-terminal truncated C5a peptides -
XX
PS Disclosure; Page 30; 84pp; English.
XX
CC The present sequence is a C-terminal truncated C5a peptide fragment. The
CC present invention relates to an antibody specific for the present
CC sequence. The C5a-antibody can be used in a therapeutic composition,
CC which is useful for treating a subject suffering from bacterial
CC infection, e.g. sinusitis, meningitis, respiratory, gastrointestinal or
CC urinary tract infections or infections in wounds. In addition, the C5a
CC antibody can be used for treating sepsis. C5a is also known as
CC anaphylatoxin.
XX
SQ Sequence 20 AA;

Query Match 89.5%; Score 102; DB 22; Length 20;
Best Local Similarity 90.0%; Pred. No. 4.1e-09;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCYDGASVNNDETCEQRAAR 20
||| |||||| ||||||||||
Db 1 CCYQGASVNNQETCEQRAAR 20

Search completed: December 9, 2003, 14:14:12
Job time : 42 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 14:12:09 ; Search time 34 Seconds
(without alignments)
151.796 Million cell updates/sec

Title: US-09-651-685A-5
Perfect score: 114
Sequence: 1 CCYDGASVNNDETCEQRAAR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 6968

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	24.6	19	11	Q9QVJ9	Q9qvj9 mus sp. mep
2	26	22.8	16	6	Q8SPT4	Q8spt4 macaca mula
3	26	22.8	17	15	Q78381	Q78381 human immun
4	26	22.8	18	6	O46665	O46665 macropus ro
5	26	22.8	19	4	Q9UCD0	Q9ucd0 homo sapien
6	26	22.8	19	11	Q9QW83	Q9qw83 rattus norv
7	25	21.9	15	4	Q9UCC0	Q9ucc0 homo sapien
8	24	21.1	18	6	Q9TT81	Q9tt81 bos taurus
9	24	21.1	20	4	Q16188	Q16188 homo sapien
10	24	21.1	20	5	P82201	P82201 bombyx mori
11	23	20.2	16	4	Q8IWX4	Q8iwx4 homo sapien
12	23	20.2	18	2	Q9R4F3	Q9r4f3 escherichia
13	23	20.2	18	2	O30888	O30888 bradyrhizob
14	23	20.2	18	6	O97668	O97668 equus cabal
15	23	20.2	19	2	Q9R5I1	Q9r5i1 brevibacter
16	23	20.2	20	4	Q9UDF5	Q9udf5 homo sapien
17	22.5	19.7	16	2	Q47605	Q47605 escherichia
18	22.5	19.7	17	5	Q8I6R5	Q8i6r5 conus imper
19	22	19.3	10	2	P96421	P96421 neisseria g
20	22	19.3	12	12	Q85666	Q85666 reovirus (t
21	22	19.3	15	11	Q9QVB6	Q9qvb6 rattus sp.
22	22	19.3	15	11	Q9QVB7	Q9qvb7 rattus sp.
23	22	19.3	17	5	Q8I6R4	Q8i6r4 conus imper
24	22	19.3	18	4	Q9UM83	Q9um83 homo sapien
25	22	19.3	18	6	Q95MX1	Q95mx1 syncerus ca
26	22	19.3	18	6	Q95MW8	Q95mw8 oryx dammah
27	22	19.3	18	6	Q95MX0	Q95mx0 tragelaphus
28	22	19.3	18	6	Q95MV5	Q95mv5 raphicerus
29	22	19.3	18	6	Q95MW7	Q95mw7 hippotragus
30	22	19.3	18	6	Q95MW1	Q95mw1 kobus leche
31	22	19.3	18	6	Q95MW3	Q95mw3 redunca ful
32	22	19.3	18	6	Q95MW5	Q95mw5 alcelaphus
33	22	19.3	18	6	Q95MW0	Q95mw0 pelea capre
34	22	19.3	18	6	Q95MV4	Q95mv4 neotragus m
35	22	19.3	18	6	Q95MW6	Q95mw6 beatragus h
36	22	19.3	18	6	Q95MW4	Q95mw4 connochaete
37	22	19.3	18	6	Q95MW2	Q95mw2 redunca red
38	22	19.3	18	6	Q95MV8	Q95mv8 litocranius
39	22	19.3	18	6	Q95MV9	Q95mv9 antidorcas
40	22	19.3	18	6	Q95MV2	Q95mv2 cephalophus
41	22	19.3	18	6	Q95MV1	Q95mv1 sylvicapra
42	22	19.3	18	6	Q95MV3	Q95mv3 oreotragus
43	22	19.3	18	6	Q95MV7	Q95mv7 raphicerus
44	22	19.3	18	6	Q95MV0	Q95mv0 ourebia our
45	22	19.3	18	6	Q95MV6	Q95mv6 raphicerus

ALIGNMENTS

RESULT 1
Q9QVJ9
ID Q9QVJ9 PRELIMINARY; PRT; 19 AA.
AC Q9QVJ9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Meprin-B peptide B1 (Fragment).
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE.
RX MEDLINE=91373354; PubMed=1894622;
RA Kounnas M.Z., Wolz R.L., Gorbea C.M., Bond J.S.;
RT "Meprin-A and -B. Cell surface endopeptidases of the mouse kidney.";
RL J. Biol. Chem. 266:17350-17357(1991).
FT NON_TER 1 1
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2157 MW; 81E66F19417E20C5 CRC64;

Query Match 24.6%; Score 28; DB 11; Length 19;
Best Local Similarity 36.4%; Pred. No. 8.9e+02;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```
Qy          3 YDGASVNNDET 13
              ::  |: ||:|
Db          1 FNQVSITNDDT 11
```

RESULT 2
Q8SPT4
ID Q8SPT4 PRELIMINARY; PRT; 16 AA.
AC Q8SPT4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chloride channel 2 (Fragment).

```
GN   CLCN2.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Norgren R.B. Jr., Zink M.A., Jia Y., Ojeda S.R., Spindel E.R.;
RT   "Construction of a targeted rhesus macaque microarray.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF489851; AAL96695.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   16 AA;  1691 MW;  50D83BF4FC6AF49E CRC64;

  Query Match             22.8%;  Score 26;  DB 6;  Length 16;
  Best Local Similarity   25.0%;  Pred. No. 1.6e+03;
  Matches    3;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 DGASVNNDETCE 15
              :|:  ::|: |:
Db          5 EGSPSDSDDKCQ 16


RESULT 3
Q78381
ID   Q78381      PRELIMINARY;      PRT;    17 AA.
AC   Q78381;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Viral sample FLPBR5F (Florida patient B), partial env cds, V5 region
DE   (Fragment).
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhang L.Q., Leigh-Brown A.J.;
RL   Submitted (APR-1992) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92271245; PubMed=1589796;
RA   Ou C.-Y., Ciesielski C.A., Myers G., Bandea C.I., Luo C.C.,
RA   Korber B.T.M., Mullins J.I., Schochetman G., Berkelman R.L.,
RA   Economou A.N., Witte J.J., Furman L.J., Satten G.A., Curran J.W.,
RA   Jaffe H.W.;
RT   "Molecular Epidemiology of HIV Transmission in a Dental Practice.";
RL   Science 256:1165-1171(1992).
DR   EMBL; M92126; AAA44496.1; -.
FT   NON_TER       1      1
FT   NON_TER      17     17
SQ   SEQUENCE   17 AA;  1708 MW;  347570D2D12CA370 CRC64;

  Query Match             22.8%;  Score 26;  DB 15;  Length 17;
  Best Local Similarity   55.6%;  Pred. No. 1.7e+03;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 GASVNNDET 13
              | : || ||
Db          2 GNNTNNTET 10


RESULT 4
O46665
ID   O46665      PRELIMINARY;      PRT;    18 AA.
AC   O46665;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Glucose-6-phosphate dehydrogenase (Fragment).
GN   G6PD.
OS   Macropus robustus robustus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX   NCBI_TaxID=35580;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97224585; PubMed=9060417;
RA   Loebel D.A., Johnston P.G.;
RT   "Analysis of the intron-exon structure of the G6PD gene of the
RT   wallaroo (Macropus robustus) by polymerase chain reaction.";
RL   Mamm. Genome 8:146-147(1997).
DR   EMBL; U53775; AAC48790.1; -.
DR   InterPro; IPR001282; G6PD.
DR   Pfam; PF02781; G6PD_C; 1.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  1992 MW;  C6D5981B528258FB CRC64;

  Query Match             22.8%;  Score 26;  DB 6;  Length 18;
  Best Local Similarity   33.3%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 ASVNNDETCEQR 17
              || |:|:  :::
Db          7 ASTNSDDVRDEK 18


RESULT 5
Q9UCD0
ID   Q9UCD0      PRELIMINARY;      PRT;    19 AA.
AC   Q9UCD0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   C215 antigen (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=94043256; PubMed=7693697;
RA   Bjork P., Jonsson U., Svedberg H., Larsson K., Lind P., Dillner J.,
RA   Hedlund G., Dohlsten M., Kalland T.;
RT   "Isolation, partial characterization, and molecular cloning of a human
RT   colon adenocarcinoma cell-surface glycoprotein recognized by the C215
RT   mouse monoclonal antibody.";
RL   J. Biol. Chem. 268:24232-24241(1993).
SQ   SEQUENCE   19 AA;  2042 MW;  54135D12119705E6 CRC64;

  Query Match             22.8%;  Score 26;  DB 4;  Length 19;
  Best Local Similarity   62.5%;  Pred. No. 1.9e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 DGASVNND 11
              :||  |||
Db          4 EGALQNND 11


RESULT 6
Q9QW83
ID   Q9QW83      PRELIMINARY;      PRT;    19 AA.
AC   Q9QW83;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   PHOSPHOPHORYN=PEPTIDE fragment 12-44 (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91291127; PubMed=2064607;
```

```
RA   Sabsay B., Stetler-Stevenson W.G., Lechner J.H., Veis A.;
RT   "Domain structure and sequence distribution in dentin phosphophoryn.";
RL   Biochem. J. 276:699-707(1991).
DR   HSSP; P00760; 1AQ7.
FT   NON_TER       1      1
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2195 MW;  744603FE729FDE0C CRC64;

  Query Match             22.8%;  Score 26;  DB 11;  Length 19;
  Best Local Similarity   44.4%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 YDGASVNND 11
              |:  ::|||
Db          7 YNSNTLNND 15


RESULT 7
Q9UCC0
ID   Q9UCC0      PRELIMINARY;      PRT;    15 AA.
AC   Q9UCC0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   Insulin-like growth factor receptor alpha subunit (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=94079885; PubMed=8257688;
RA   Kasuya J., Paz I.B., Maddux B.A., Goldfine I.D., Hefta S.A.,
RA   Fujita-Yamaguchi Y.;
RT   "Characterization of human placental insulin-like growth factor-
RT   I/insulin hybrid receptors by protein microsequencing and
RT   purification.";
RL   Biochemistry 32:13531-13536(1993).
SQ   SEQUENCE   15 AA;  1721 MW;  98BC151D6D81784B CRC64;

  Query Match             21.9%;  Score 25;  DB 4;  Length 15;
  Best Local Similarity   40.0%;  Pred. No. 2.2e+03;
  Matches    4;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNND 11
              |   |  : ||
Db          3 CGPGIDIRND 12


RESULT 8
Q9TT81
ID   Q9TT81      PRELIMINARY;      PRT;    18 AA.
AC   Q9TT81;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Wilm's tumor protein 1 (Fragment).
GN   WT1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21015404; PubMed=11130975;
RA   Brouillette J.A., Andrew J.R., Venta P.J.;
RT   "Estimate of nucleotide diversity in dogs with a pool-and-sequence
RT   method.";
RL   Mamm. Genome 11:1079-1086(2000).
DR   EMBL; AF201738; AAF19824.1; -.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2196 MW;  1EE8DAD8E072C0BF CRC64;

  Query Match             21.1%;  Score 24;  DB 6;  Length 18;
  Best Local Similarity   33.3%;  Pred. No. 4e+03;
  Matches    3;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         12 ETCEQRAAR 20
              :|:::: :|
Db          7 KTCQRKFSR 15


RESULT 9
Q16188
ID   Q16188      PRELIMINARY;      PRT;    20 AA.
AC   Q16188;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Adenosine deaminase protein (Fragment).
GN   ADENOSINE DEAMINASE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94327968; PubMed=8051429;
RA   Shovlin C.L., Simmonds H.A., Fairbanks L.D., Deacock S.J.,
RA   Hughes J.M., Lechler R.I., Webster A.D., Sun X.M., Webb J.C.,
RA   Soutar A.K.;
RT   "Adult onset immunodeficiency caused by inherited adenosine deaminase
RT   deficiency.";
RL   J. Immunol. 153:2331-2339(1994).
DR   EMBL; S72469; AAD14102.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   20 AA;  2175 MW;  4EABA6143F739B4C CRC64;

  Query Match             21.1%;  Score 24;  DB 4;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 4.5e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         13 TCEQRAAR 20
              :|| | :|
Db         12 SCEVRCSR 19


RESULT 10
P82201
ID   P82201      PRELIMINARY;      PRT;    20 AA.
AC   P82201;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   Unknown protein from 2D-page (Fragment).
OS   Bombyx mori (Silk moth).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC   Bombycidae; Bombyx.
OX   NCBI_TaxID=7091;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=XINHANG X KEMING; TISSUE=Body wall, and Fat body;
RX   MEDLINE=21177481; PubMed=11280994;
RA   Zhong B.X.;
RT   "Protein database for several tissues derived from five instar of
RT   silkworm.";
RL   I Chuan Hsueh Pao 28:217-224(2001).
CC   -!- SIMILARITY: TO THE N-TERMINAL OF TROPOMYOSINS.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2215 MW;  A3C1EFD3F1350767 CRC64;

  Query Match             21.1%;  Score 24;  DB 5;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 4.5e+03;
```

```
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 CEQRA 18
              |||:|
Db          5 CEQQA 9


RESULT 11
Q8IWX4
ID   Q8IWX4      PRELIMINARY;      PRT;    16 AA.
AC   Q8IWX4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Transcription factor AP-2 gamma (Fragment).
GN   TFAP2G.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22378671; PubMed=12490322;
RA   Li M., Wang Y., Yu Y., Nishizawa M., Nakajima T., Ito S., Kannan P.;
RT   "The human transcription factor activation protein-2 gamma (AP-
RT   2gamma): gene structure, promoter, and expression in mammary carcinoma
RT   cell lines.";
RL   Gene 301:43-51(2002).
DR   EMBL; AF544986; AAN37906.1; -.
FT   NON_TER      16      16
SQ   SEQUENCE   16 AA;  2016 MW;  F88CAEF56EEFC601 CRC64;

  Query Match             20.2%;  Score 23;  DB 4;  Length 16;
  Best Local Similarity   44.4%;  Pred. No. 5.2e+03;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 SVNNDETCE 15
              :|  :| ||
Db          8 NVKYEEDCE 16


RESULT 12
Q9R4F3
ID   Q9R4F3      PRELIMINARY;      PRT;    18 AA.
AC   Q9R4F3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   20 kDa fimbrial adhesin subunit (Fragment).
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=96110952; PubMed=8557360;
RA   Bertin Y., Girardeau J.P., Darfeuille-Michaud A., Contrepois M.;
RT   "Characterization of 20K fimbria, a new adhesin of septicemic and
RT   diarrhea-associated Escherichia coli strains, that belongs to a family
RT   of adhesins with N-acetyl-D-glucosamine recognition.";
RL   Infect. Immun. 64:332-342(1996).
SQ   SEQUENCE   18 AA;  1990 MW;  79AF96A13476DF81 CRC64;

  Query Match             20.2%;  Score 23;  DB 2;  Length 18;
  Best Local Similarity   35.7%;  Pred. No. 5.9e+03;
  Matches    5;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 YDGASVNNDETCEQ 16
              |||    | :  :|
Db          1 YDGTINFNGKVVDQ 14


RESULT 13
O30888
ID   O30888      PRELIMINARY;      PRT;    18 AA.
AC   O30888;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Nodulation protein (Fragment).
GN   NODA.
OS   Bradyrhizobium sp.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Bradyrhizobium.
OX   NCBI_TaxID=376;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20020861; PubMed=10553291;
RA   Sterner J.P., Parker M.A.;
RT   "Diversity and relationships of Bradyrhizobia from Amphicarpaea
RT   bracteata based on partial nod and ribosomal sequences.";
RL   Syst. Appl. Microbiol. 22:387-392(1999).
DR   EMBL; AF020680; AAB71696.1; -.
FT   NON_TER      18      18
SQ   SEQUENCE   18 AA;  1943 MW;  C8C595DCD14D7A5F CRC64;

  Query Match             20.2%;  Score 23;  DB 2;  Length 18;
  Best Local Similarity   28.6%;  Pred. No. 5.9e+03;
  Matches    4;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          7 SVNNDETCEQRAAR 20
              ::    : | |:||
Db          2 NIPGSRSAESRSAR 15


RESULT 14
O97668
ID   O97668      PRELIMINARY;      PRT;    18 AA.
AC   O97668;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Lactoferrin (Fragment).
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Brandon R.B., Giffard J.M., Bell T.K.;
RT   "Isolation of Equine Lactoferrin Gene.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF102025; AAC77463.1; -.
DR   HSSP; O77698; 1CE2.
DR   InterPro; IPR001156; Transferrin.
DR   Pfam; PF00405; transferrin; 1.
FT   NON_TER       1       1
FT   NON_TER      18      18
SQ   SEQUENCE   18 AA;  1936 MW;  D8FBFA226AA6DDD9 CRC64;

  Query Match             20.2%;  Score 23;  DB 6;  Length 18;
  Best Local Similarity   66.7%;  Pred. No. 5.9e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CCYDGA 6
              |  |||
Db          1 CLQDGA 6


RESULT 15
Q9R5I1
ID   Q9R5I1      PRELIMINARY;      PRT;    19 AA.
AC   Q9R5I1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

DE Nitrile hydratase alpha subunit (Fragment).
OS Brevibacterium.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Brevibacteriaceae.
OX NCBI_TaxID=1696;
RN [1]
RP *SEQUENCE*.
RX MEDLINE=92407758; PubMed=1527703;
RA Duran R., Chion C.K., Bigey F., Arnaud A., Galzy P.;
RT "The N-terminal amino acid sequences of Brevibacterium sp. R312
RT nitrile hydratase.";
RL J. Basic Microbiol. 32:13-19(1992).
SQ SEQUENCE 19 AA; 1923 MW; 84726D1A1282FB63 CRC64;

Query Match 20.2%; Score 23; DB 2; Length 19;
Best Local Similarity 50.0%; Pred. No. 6.3e+03;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          7 SVNNDETCEQRA 18
              ||  | | |  |
Db          1 SVTIDHTTENAA 12
```

Search completed: December 9, 2003, 14:15:22
Job time : 35 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 14:09:04 ; Search time 11 Seconds
(without alignments)
85.503 Million cell updates/sec

Title: US-09-651-685A-5
Perfect score: 114
Sequence: 1 CCYDGASVNNDETCEQRAAR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 1220

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	31.5	27.6	16	1	CXA1_CONEP	P56638	conus episc
2	29.5	25.9	18	1	CXA1_CONER	P50982	conus ermin
3	29	25.4	15	1	ITRB_ALBJU	P24927	albizzia ju
4	26	22.8	13	1	CXAA_CONST	P28878	conus stria
5	25	21.9	16	1	CXA1_CONAL	P56639	conus aulic
6	25	21.9	16	1	CXAA_CONPE	P50984	conus penna
7	25	21.9	18	1	HSTB_ECOLI	P01560	escherichia
8	25	21.9	19	1	LCRP_PETMA	Q10996	petromyzon
9	24	21.1	14	1	SCK3_LEIQU	P45661	leiurus qui
10	24	21.1	20	1	SB60_MAIZE	P82868	zea mays (m
11	23	20.2	14	1	CXA1_CONMA	P01521	conus magus
12	22	19.3	12	1	CXA1_CONIM	P50983	conus imper
13	21	18.4	11	1	NXSN_PSETE	P59072	pseudonaja
14	21	18.4	15	1	NXSO_PSETE	P59073	pseudonaja
15	21	18.4	15	1	UN01_PINPS	P81106	pinus pinas
16	21	18.4	16	1	CXA3_CONAL	P56641	conus aulic
17	21	18.4	16	1	CXAB_CONPE	P50985	conus penna
18	21	18.4	18	1	PPM2_LIMPO	P14216	limulus pol
19	21	18.4	20	1	SC18_MESMA	P58570	mesobuthus
20	20	17.5	10	1	RRPL_PHODV	P35946	phocine dis
21	20	17.5	11	1	CX5A_CONAL	P58848	conus aulic
22	20	17.5	11	1	CX5B_CONAL	P58849	conus aulic
23	20	17.5	13	1	CXA2_CONGE	P01520	conus geogr
24	20	17.5	14	1	CX1A_CONBE	P58623	conus betul
25	20	17.5	14	1	CXA1_CONCN	P56973	conus conso
26	20	17.5	15	1	CX1B_CONBE	P58624	conus betul
27	20	17.5	15	1	OBPA_MAMBR	P81285	mamestra br
28	20	17.5	15	1	TERM_BPM2	P19897	bacteriopha
29	20	17.5	16	1	FOR1_MYRGU	P81438	myrmecia gu
30	20	17.5	17	1	TAC1_TACGI	P23684	tachypleus
31	20	17.5	17	1	TAC3_TACGI	P18252	tachypleus
32	20	17.5	18	1	PPM1_LIMPO	P14215	limulus pol
33	20	17.5	19	1	HI70_RAT	P21794	rattus norv
34	19	16.7	9	1	OXYA_SCYCA	P42996	scyliorhinu
35	19	16.7	9	1	OXYA_SQUAC	P42999	squalus aca
36	19	16.7	10	1	COXA_ONCMY	P80328	oncorhynchu
37	19	16.7	10	1	PPCK_FASHE	P80525	fasciola he
38	19	16.7	10	1	UH05_RAT	P56573	rattus norv
39	19	16.7	15	1	GUAN_DIDMA	P55936	didelphis m
40	19	16.7	16	1	ARCD_PSEPU	P41147	pseudomonas
41	19	16.7	17	1	CXMA_CONPE	P58926	conus penna
42	19	16.7	17	1	CXMB_CONPE	P58927	conus penna
43	19	16.7	17	1	PATS_ANASP	O52748	anabaena sp
44	19	16.7	18	1	RL24_PROVU	P20032	proteus vul
45	19	16.7	18	1	UC03_MAIZE	P80609	zea mays (m

ALIGNMENTS

RESULT 1
CXA1_CONEP
ID CXA1_CONEP STANDARD; PRT; 16 AA.
AC P56638;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-conotoxin EpI.
OS Conus episcopatus (Bishop's cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=88764;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS).
RX MEDLINE=98376423; PubMed=9708977;
RA Hu S.H., Loughnan M., Miller R., Weeks C.M., Blessing R.H.,
RA Alewood P.F., Lewis R.J., Martin J.L.;
RT "The 1.1-A resolution crystal structure of [Tyr15]EpI, a novel
RT alpha-conotoxin from Conus episcopatus, solved by direct methods.";
RL Biochemistry 37:11425-11433(1998).
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM. THIS PEPTIDE BLOCKS MAMMALIAN NICOTINIC
CC ACETYLCHOLINE RECEPTORS COMPOSED OF ALPHA-3/BETA-2 AND ALPHA-
CC 3/BETA-4 SUBUNITS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC FAMILY.
DR PIR; A59042; A59042.
DR PDB; 1A0M; 13-JAN-99.
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW Acetylcholine receptor inhibitor; Amidation;
KW Sulfation; 3D-structure.
FT DISULFID 2 8
FT DISULFID 3 16
FT MOD_RES 15 15 SULFATION.
FT MOD_RES 16 16 AMIDATION.
FT HELIX 2 4
FT HELIX 6 11
FT TURN 13 16
SQ SEQUENCE 16 AA; 1792 MW; C63385F376C99B4C CRC64;

Query Match 27.6%; Score 31.5; DB 1; Length 16;
Best Local Similarity 40.0%; Pred. No. 43;
Matches 6; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

```
Qy          1 CCYD-GASVNNDETC 14
              || |   ::|| : |
Db          2 CCSDPRCNMNNPDYC 16
```

RESULT 2
CXA1_CONER
ID CXA1_CONER STANDARD; PRT; 18 AA.

```
AC   P50982;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alpha-conotoxin EI.
OS   Conus ermineus (Atlantic fish-hunting cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=55423;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=96062516; PubMed=7578057;
RA   Martinez J.S., Olivera B.M., Gray W.R., Craig A.G., Groebe D.R.,
RA   Abramson S.N., McIntosh J.M.;
RT   "Alpha-conotoxin EI, a new nicotinic acetylcholine receptor
RT   antagonist with novel selectivity.";
RL   Biochemistry 34:14519-14526(1995).
CC   -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC       BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC       INHIBIT THEM.
CC   -!- SUBUNIT: Binds nicotinic acetylcholine receptor.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC   -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC       FAMILY.
DR   PIR; A58589; A58589.
KW   Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW   Acetylcholine receptor inhibitor; Amidation; Hydroxylation.
FT   DISULFID      4     10
FT   DISULFID      5     18
FT   MOD_RES       3      3       HYDROXYLATION.
FT   MOD_RES      18     18       AMIDATION.
SQ   SEQUENCE   18 AA;  2082 MW;  60A61A6C427A6B5E CRC64;

  Query Match             25.9%;  Score 29.5;  DB 1;  Length 18;
  Best Local Similarity   33.3%;  Pred. No. 1e+02;
  Matches    5;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 CCY-DGASVNNDETC 14
              |||    :::| : |
Db          4 CCYHPTCNMSNPQIC 18


RESULT 3
ITRB_ALBJU
ID   ITRB_ALBJU     STANDARD;      PRT;    15 AA.
AC   P24927;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Trypsin inhibitor B chain (Fragment).
OS   Albizzia julibrissin (Silk tree).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Mimosoideae; Ingeae; Albizia.
OX   NCBI_TaxID=3813;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Seed;
RX   MEDLINE=80115605; PubMed=528539;
RA   Odani S., Ono T., Ikenaka T.;
RT   "Proteinase inhibitors from a mimosoideae legume, Albizzia
RT   julibrissin. Homologues of soybean trypsin inhibitor (Kunitz).";
RL   J. Biochem. 86:1795-1805(1979).
CC   -!- FUNCTION: INHIBITS TRYPSIN AND ALPHA-CHYMOTRYPSIN.
CC   -!- SUBUNIT: HETERODIMER OF AN "A" AND A "B" CHAIN LINKED BY A
CC       DISULFIDE BOND.
CC   -!- SIMILARITY: BELONGS TO THE LEGUMINOUS KUNITZ-TYPE INHIBITOR
CC       FAMILY.
KW   Serine protease inhibitor.
FT   NON_TER       15     15
SQ   SEQUENCE   15 AA;  1705 MW;  53165F7E9C45B4D0 CRC64;

  Query Match             25.4%;  Score 29;  DB 1;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 1e+02;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 CYDGASVNNDE 12
              | |  |:::||
Db          5 CKDLGSIDDDE 15


RESULT 4
CXAA_CONST
ID   CXAA_CONST     STANDARD;      PRT;    13 AA.
AC   P28878;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alpha-conotoxin SIA (S1A).
OS   Conus striatus (Striated cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=6493;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=91369955; PubMed=1892838;
RA   Myers R.A., Zafarella G.C., Gray W.R., Abbot J., Cruz L.J.,
RA   Olivera B.M.;
RT   "Alpha-conotoxins, small peptide probes of nicotinic acetylcholine
RT   receptors.";
RL   Biochemistry 30:9370-9377(1991).
CC   -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC       BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC       INHIBIT THEM.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC   -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC       FAMILY.
DR   PIR; A40312; NTKNAS.
KW   Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW   Acetylcholine receptor inhibitor; Amidation.
FT   DISULFID      2      7
FT   DISULFID      3     13
FT   MOD_RES      13     13       AMIDATION.
SQ   SEQUENCE   13 AA;  1461 MW;  DEF1931982457EBD CRC64;

  Query Match             22.8%;  Score 26;  DB 1;  Length 13;
  Best Local Similarity   36.4%;  Pred. No. 2.6e+02;
  Matches    4;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNND 11
              ||:     | |
Db          2 CCHPACGKNFD 12


RESULT 5
CXA1_CONAL
ID   CXA1_CONAL     STANDARD;      PRT;    16 AA.
AC   P56639;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alpha-conotoxin AuIA.
OS   Conus aulicus (Court cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=89437;
RN   [1]
RP   SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC   TISSUE=Venom;
```

```
RX   MEDLINE=99003392; PubMed=9786965;
RA   Luo S., Kulak J.M., Cartier G.E., Jacobsen R.B., Yoshikami D.,
RA   Olivera B.M., McIntosh J.M.;
RT   "Alpha-conotoxin AuIB selectively blocks alpha3 beta4 nicotinic
RT   acetylcholine receptors and nicotine-evoked norepinephrine release.";
RL   J. Neurosci. 18:8571-8579(1998).
CC   -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC       BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC       INHIBIT THEM. THIS PEPTIDE BLOCKS MAMMALIAN NICOTINIC
CC       ACETYLCHOLINE RECEPTORS COMPOSED OF ALPHA-3/BETA-4 SUBUNITS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC   -!- MASS SPECTROMETRY: MW=1725.6; METHOD=Electrospray.
CC   -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC       FAMILY.
DR   PIR; A59045; A59045.
DR   HSSP; P50984; 1PEN.
KW   Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW   Acetylcholine receptor inhibitor; Amidation.
FT   DISULFID      2      8
FT   DISULFID      3     16
FT   MOD_RES      16     16       AMIDATION.
SQ   SEQUENCE   16 AA;  1731 MW;  1E310FEB8FDC7001 CRC64;

  Query Match             21.9%;  Score 25;  DB 1;  Length 16;
  Best Local Similarity   28.6%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    1;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETC 14
              | |       | : |
Db          3 CSYPPCPATNSDYC 16


RESULT 6
CXAA_CONPE
ID   CXAA_CONPE     STANDARD;      PRT;    16 AA.
AC   P50984;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alpha-conotoxin PnIA.
OS   Conus pennaceus (Feathered cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=37335;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=94347719; PubMed=8068627;
RA   Fainzilber M., Hasson A., Oren R., Burlingame A.L., Gordon D.,
RA   Spira M.E., Zlotkin E.;
RT   "New mollusc-specific alpha-conotoxins block Aplysia neuronal
RT   acetylcholine receptors.";
RL   Biochemistry 33:9523-9529(1994).
RN   [2]
RP   SULFATION OF TYR-15.
RX   MEDLINE=99242956; PubMed=10226369;
RA   Wolfender J.L., Chu F., Ball H., Wolfender F., Fainzilber M.,
RA   Baldwin M.A., Burlingame A.L.;
RT   "Identification of tyrosine sulfation in Conus pennaceus conotoxins
RT   alpha-PnIA and alpha-PnIB: further investigation of labile sulfo- and
RT   phosphopeptides by electrospray, matrix-assisted laser
RT   desorption/ionization (MALDI) and atmospheric pressure MALDI mass
RT   spectrometry.";
RL   J. Mass Spectrom. 34:447-454(1999).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS).
RX   MEDLINE=96311277; PubMed=8740364;
RA   Hu S.-H., Gehrmann J., Guddat L.W., Alewood P.F., Craik D.J.,
RA   Martin J.L.;
RT   "The 1.1 A crystal structure of the neuronal acetylcholine receptor
RT   antagonist, alpha-conotoxin PnIA from Conus pennaceus.";
RL   Structure 4:417-423(1996).
CC   -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC       BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC       INHIBIT THEM. IN CONTRAST TO OTHER ALPHA-CONOTOXINS, WHICH ARE
CC       SELECTIVE FOR VERTEBRATE SKELETAL MUSCLE NACHR, THE CONUS
CC       PENNACEUS ALPHA-CONOTOXINS BLOCK NACHR IN MOLLUSKS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC   -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC       FAMILY.
DR   PIR; A54877; A54877.
DR   PDB; 1PEN; 21-APR-97.
KW   Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW   Acetylcholine receptor inhibitor; Amidation; Sulfation; 3D-structure.
FT   DISULFID      2      8
FT   DISULFID      3     16
FT   MOD_RES      15     15       SULFATION.
FT   MOD_RES      16     16       AMIDATION.
FT   HELIX         2      4
FT   HELIX         6     11
FT   TURN         13     16
SQ   SEQUENCE   16 AA;  1628 MW;  05310FF95EC99005 CRC64;

  Query Match             21.9%;  Score 25;  DB 1;  Length 16;
  Best Local Similarity   28.6%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETC 14
              |     : || : |
Db          3 CSLPPCAANNPDYC 16


RESULT 7
HSTB_ECOLI
ID   HSTB_ECOLI     STANDARD;      PRT;    18 AA.
AC   P01560;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Heat-stable enterotoxin ST-2 (ST-B).
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=O42:K86:H37 / 18D;
RX   MEDLINE=81264141; PubMed=7021541;
RA   Chan S.-K., Giannella R.A.;
RT   "Amino acid sequence of heat-stable enterotoxin produced by
RT   Escherichia coli pathogenic for man.";
RL   J. Biol. Chem. 256:7744-7746(1981).
RN   [2]
RP   DISULFIDE BONDS.
RX   MEDLINE=87191003; PubMed=3552731;
RA   Shimonishi Y., Hidaka Y., Koizumi M., Hane M., Aimoto S., Takeda T.,
RA   Miwatani T., Takeda Y.;
RT   "Mode of disulfide bond formation of a heat-stable enterotoxin (STh)
RT   produced by a human strain of enterotoxigenic Escherichia coli.";
RL   FEBS Lett. 215:165-170(1987).
CC   -!- FUNCTION: TOXIN WHICH ACTIVATES THE PARTICULATE FORM OF GUANYLATE
CC       CYCLASE AND INCREASES CYCLIC GMP LEVELS WITHIN THE HOST
CC       INTESTINAL EPITHELIAL CELLS.
CC   -!- DISEASE: BOTH HEAT-STABLE AND HEAT-LABILE ENTEROTOXINS ARE
CC       PRODUCED BY PATHOGENIC STRAINS OF E.COLI AND AFFECT THE DIGESTIVE
CC       TRACT OF MAMMALS.
CC   -!- SIMILARITY: BELONGS TO THE HEAT-STABLE ENTEROTOXIN FAMILY.
DR   PIR; A01823; QHEC2.
DR   HSSP; P01559; 1ETN.
DR   InterPro; IPR001489; Enterotoxin_HS.
DR   Pfam; PF02048; Enterotoxin_HS; 1.
DR   PROSITE; PS00273; ENTEROTOXIN_H_STABLE; 1.
KW   Toxin; Enterotoxin.
```

```
FT   DISULFID       5      10
FT   DISULFID       6      14
FT   DISULFID       9      17
SQ   SEQUENCE   18 AA;  1978 MW;  D0C975F49D600650 CRC64;

  Query Match             21.9%;  Score 25;  DB 1;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+02;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCY 3
              |||
Db          9 CCY 11


RESULT 8
LCRP_PETMA
ID   LCRP_PETMA     STANDARD;      PRT;    19 AA.
AC   Q10996;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Corticostatin-related protein LCRP.
OS   Petromyzon marinus (Sea lamprey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC   Petromyzontiformes; Petromyzontidae; Petromyzon.
OX   NCBI_TaxID=7757;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Skin;
RX   MEDLINE=96321324; PubMed=8759287;
RA   Conlon J.M., Sower S.A.;
RT   "Isolation of a peptide structurally related to mammalian
RT   corticostatins from the lamprey Petromyzon marinus.";
RL   Comp. Biochem. Physiol. 114B:133-137(1996).
CC   -!- FUNCTION: MAY HAVE MICROBICIDAL ACTIVITIES. MAY INHIBIT
CC       CORTICOTROPIN (ACTH) STIMULATED STEROIDOGENESIS AND THE MICROBIAL
CC       ACTIONS OF THE CORTICOSTATINS.
CC   -!- MASS SPECTROMETRY: MW=2201; MW_ERR=0.4; METHOD=Electrospray.
CC   -!- SIMILARITY: BELONGS TO THE CORTICOSTATIN/DEFENSIN FAMILY.
KW   Defensin; Antibiotic.
FT   DISULFID       1      18       BY SIMILARITY.
FT   DISULFID       3       9       BY SIMILARITY.
FT   DISULFID       8      17       BY SIMILARITY.
SQ   SEQUENCE   19 AA;  2209 MW;  8D9CEDC71A199AE5 CRC64;

  Query Match             21.9%;  Score 25;  DB 1;  Length 19;
  Best Local Similarity   50.0%;  Pred. No. 5.7e+02;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CCYDGASV 8
              ||  | :|
Db          8 CCVRGLNV 15


RESULT 9
SCK3_LEIQU
ID   SCK3_LEIQU     STANDARD;      PRT;    14 AA.
AC   P45661;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leiurutoxin III (Fragment).
OS   Leiurus quinquestriatus quinquestriatus (Egyptian scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Leiurus.
OX   NCBI_TaxID=6885;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=93075256; PubMed=1280139;
RA   Valdivia H.H., Martin B.M., Escobar L., Possani L.D.;
RT   "Noxiustoxin and leiurutoxin III, two homologous peptide toxins with
RT   binding properties to synaptosomal membrane K+ channels.";
RL   Biochem. Int. 27:953-962(1992).
CC   -!- FUNCTION: Blocker of potassium channels.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE SHORT SCORPION TOXIN FAMILY. POTASSIUM
CC       CHANNEL INHIBITORS SUBFAMILY.
DR   PIR; A48389; A48389.
DR   InterPro; IPR001947; Scorpion_toxinS.
DR   PROSITE; PS01138; SCORP_SHORT_TOXIN; PARTIAL.
KW   Toxin; Neurotoxin; Ionic channel inhibitor;
KW   Potassium channel inhibitor.
FT   NON_TER       14      14
SQ   SEQUENCE   14 AA;  1588 MW;  83C67CCBD691205E CRC64;

  Query Match             21.1%;  Score 24;  DB 1;  Length 14;
  Best Local Similarity   35.7%;  Pred. No. 5.9e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    6;  Gaps    1;

Qy          2 CYDGASVNNDETCE 15
              ||| :       ||
Db          7 CYDSSQ------CE 14


RESULT 10
SB60_MAIZE
ID   SB60_MAIZE     STANDARD;      PRT;    20 AA.
AC   P82868;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Putative 60 kDa spermidine-binding protein (Fragment).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE, AND SUBCELLULAR LOCATION.
RC   STRAIN=cv. Cecilia; TISSUE=Coleoptile;
RX   MEDLINE=21948208; PubMed=11950979;
RA   Tassoni A., Napier R.M., Franceschetti M., Venis M.A., Bagni N.;
RT   "Spermidine-binding proteins. Purification and expression analysis in
RT   maize.";
RL   Plant Physiol. 128:1303-1312(2002).
CC   -!- FUNCTION: May have spermidine-binding activity.
CC   -!- SUBUNIT: Dimer of 18 kDa and 60 kDa subunit (Probable).
CC   -!- SUBCELLULAR LOCATION: Microsomal membrane.
CC   -!- MISCELLANEOUS: On the 2D-gel its MW is: 60 kDa.
DR   GO; GO:0005792; C:microsome; NAS.
KW   Membrane; Microsome.
FT   NON_TER       20      20
SQ   SEQUENCE   20 AA;  2093 MW;  7FD730C00391579A CRC64;

  Query Match             21.1%;  Score 24;  DB 1;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 8.7e+02;
  Matches    6;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 ASVNNDETCEQRAAR 20
              | |    || : | |:
Db          4 AVVEPPETSQNRIAK 18


RESULT 11
CXA1_CONMA
ID   CXA1_CONMA     STANDARD;      PRT;    14 AA.
AC   P01521;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alpha-conotoxin MI (M1).
OS   Conus magus (Magus cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
```

```
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=6492;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=83073458; PubMed=7149738;
RA   McIntosh J.M., Cruz L.J., Hunkapiller M.W., Gray W.R., Olivera B.M.;
RT   "Isolation and structure of a peptide toxin from the marine snail
RT   Conus magus.";
RL   Arch. Biochem. Biophys. 218:329-334(1982).
RN   [2]
RP   DISULFIDE BONDS.
RX   MEDLINE=84032400; PubMed=6630187;
RA   Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;
RT   "Conotoxin MI. Disulfide bonding and conformational states.";
RL   J. Biol. Chem. 258:12247-12251(1983).
CC   -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC       BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC       INHIBIT THEM.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC   -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC       FAMILY.
DR   PIR; A01784; NTKN1M.
DR   HSSP; P56973; 1B45.
KW   Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW   Acetylcholine receptor inhibitor; Amidation.
FT   DISULFID      3      8
FT   DISULFID      4     14
FT   MOD_RES      14     14       AMIDATION.
SQ   SEQUENCE   14 AA;  1499 MW;  DEEE91898BF5E5BD CRC64;

  Query Match             20.2%;  Score 23;  DB 1;  Length 14;
  Best Local Similarity   33.3%;  Pred. No. 8.5e+02;
  Matches    3;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVN 9
              ||:     |
Db          3 CCHPACGKN 11


RESULT 12
CXA1_CONIM
ID   CXA1_CONIM     STANDARD;      PRT;    12 AA.
AC   P50983;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alpha-conotoxin ImI.
OS   Conus imperialis (Imperial cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=35631;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Venom;
RX   MEDLINE=94266889; PubMed=8206995;
RA   McIntosh J.M., Yoshikami D., Mahe E., Nielsen D.B., Rivier J.E.,
RA   Gray W.R., Olivera B.M.;
RT   "A nicotinic acetylcholine receptor ligand of unique specificity,
RT   alpha-conotoxin ImI.";
RL   J. Biol. Chem. 269:16733-16739(1994).
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=95379776; PubMed=7651351;
RA   Johnson D.S., Martinez J., Elgoyhen A.B., Heinemann S.F.,
RA   McIntosh J.M.;
RT   "Alpha-conotoxin ImI exhibits subtype-specific nicotinic
RT   acetylcholine receptor blockade: preferential inhibition of homomeric
RT   alpha 7 and alpha 9 receptors.";
RL   Mol. Pharmacol. 48:194-199(1995).
RN   [3]
RP   STRUCTURE BY NMR.
RX   MEDLINE=99212205; PubMed=10194298;
RA   Rogers J.P., Luginbuehl P., Shen G.S., McCabe R.T., Stevens R.C.,
RA   Wemmer D.E.;
RT   "NMR solution structure of alpha-conotoxin ImI and comparison to
RT   other conotoxins specific for neuronal nicotinic acetylcholine
RT   receptors.";
RL   Biochemistry 38:3874-3882(1999).
RN   [4]
RP   STRUCTURE BY NMR.
RX   MEDLINE=99280313; PubMed=10350614;
RA   Gouda H., Hirono S.;
RT   "Solution structure of alpha-conotoxin ImI determined by
RT   two-dimensional NMR spectroscopy.";
RL   Biochim. Biophys. Acta 1431:384-394(1999).
RN   [5]
RP   STRUCTURE BY NMR.
RX   MEDLINE=99158061; PubMed=10050774;
RA   Maslennikov I.V., Shenkarev Z.O., Zhmak M.N., Ivanov V.T.,
RA   Methfessel C., Tsetlin V.I., Arseniev A.S.;
RT   "NMR spatial structure of alpha-conotoxin ImI reveals a common
RT   scaffold in snail and snake toxins recognizing neuronal nicotinic
RT   acetylcholine receptors.";
RL   FEBS Lett. 444:275-280(1999).
RN   [6]
RP   STRUCTURE BY NMR.
RX   MEDLINE=99358772; PubMed=10431825;
RA   Lamthanh H., Jegou-Matheron C., Servent D., Menez A., Lancelin J.-M.;
RT   "Minimal conformation of the alpha-conotoxin ImI for the alpha7
RT   neuronal nicotinic acetylcholine receptor recognition: correlated CD,
RT   NMR and binding studies.";
RL   FEBS Lett. 454:293-298(1999).
RN   [7]
RP   STRUCTURE BY NMR.
RX   MEDLINE=99324017; PubMed=10395477;
RA   Gehrmann J., Daly N.L., Alewood P.F., Craik D.J.;
RT   "Solution structure of alpha-conotoxin ImI by 1H nuclear magnetic
RT   resonance.";
RL   J. Med. Chem. 42:2364-2372(1999).
RN   [8]
RP   MUTAGENESIS OF ASP-5; ARG-7 AND ARG-11, AND STRUCTURE BY NMR OF THESE
RP   THREE MUTANTS.
RX   MEDLINE=20574023; PubMed=11124036;
RA   Rogers J.P., Luginbuhl P., Pemberton K., Harty P., Wemmer D.E.,
RA   Stevens R.C.;
RT   "Structure-activity relationships in a peptidic alpha7 nicotinic
RT   acetylcholine receptor antagonist.";
RL   J. Mol. Biol. 304:911-926(2000).
CC   -!- FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they
CC       bind to the nicotinic acetylcholine receptors (nAChR) and thus
CC       inhibit them. It is highly active against the neuromuscular
CC       receptor in frog but not in mice. In contrast, it induces seizures
CC       when injected centrally in mice and rats. It targets neuronal
CC       nAChRs in mammals. It blocks homomeric alpha-7 nicotinic receptors
CC       with the highest apparent affinity and homomeric alpha-9 receptors
CC       with 8-fold lower affinity. It has no effect on receptors composed
CC       of alpha-2/beta-2, alpha-3/beta-2, alpha-4/beta-2, alpha-2/beta-4,
CC       alpha-3/beta-4, or alpha-4/beta-4 subunit combinations.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC   -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
CC       FAMILY.
DR   PIR; A53709; A53709.
DR   PDB; 1IM1; 15-JUN-99.
DR   PDB; 1IMI; 23-APR-99.
DR   PDB; 1CNL; 27-MAY-99.
DR   PDB; 1E74; 27-DEC-00.
DR   PDB; 1E75; 27-DEC-00.
DR   PDB; 1E76; 27-DEC-00.
DR   PDB; 1G2G; 08-NOV-00.
KW   Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW   Acetylcholine receptor inhibitor; Amidation; 3D-structure.
FT   DISULFID      2      8
FT   DISULFID      3     12
```

```
FT   MOD_RES       12      12       AMIDATION.
FT   MUTAGEN        5       5       D->L: REDUCTION OF TOXICITY.
FT   MUTAGEN        7       7       R->L: REDUCTION OF TOXICITY.
FT   MUTAGEN       11      11       R->E: NO LOSS OF ACTIVITY.
FT   HELIX          2       4
FT   HELIX          6       8
FT   TURN          10      12
SQ   SEQUENCE   12 AA;  1357 MW;  9C29CEA545A4176A CRC64;

  Query Match             19.3%;  Score 22;  DB 1;  Length 12;
  Best Local Similarity   75.0%;  Pred. No. 1e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CCYD 4
              || |
Db          2 CCSD 5


RESULT 13
NXSN_PSETE
ID   NXSN_PSETE     STANDARD;      PRT;    11 AA.
AC   P59072;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Short neurotoxin N1 (Alpha neurotoxin) (Fragment).
OS   Pseudonaja textilis (Eastern brown snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Acanthophiinae; Pseudonaja.
OX   NCBI_TaxID=8673;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Venom;
RX   MEDLINE=99449602; PubMed=10518793;
RA   Gong N.L., Armugam A., Jeyaseelan K.;
RT   "Postsynaptic short-chain neurotoxins from Pseudonaja textilis: cDNA
RT   cloning, expression and protein characterization.";
RL   Eur. J. Biochem. 265:982-989(1999).
CC   -!- FUNCTION: Lethal neurotoxin, binds and inhibits nicotinic
CC       acetylcholine receptors (nAChR).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- MASS SPECTROMETRY: MW=6236; METHOD=Electrospray.
CC   -!- MISCELLANEOUS: LD(50) is 0.84 mg/kg by intravenous injection.
CC   -!- SIMILARITY: Belongs to the snake toxin family.
DR   InterPro; IPR003571; Snake_toxin.
DR   PROSITE; PS00272; SNAKE_TOXIN; PARTIAL.
KW   Toxin; Neurotoxin; Postsynaptic neurotoxin;
KW   Acetylcholine receptor inhibitor; Multigene family.
FT   UNSURE         3       3
FT   NON_TER       11      11
SQ   SEQUENCE   11 AA;  1319 MW;  0D1EF0C81B58732B CRC64;

  Query Match             18.4%;  Score 21;  DB 1;  Length 11;
  Best Local Similarity   75.0%;  Pred. No. 1.4e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CYDG 5
              || |
Db          3 CYKG 6


RESULT 14
NXSO_PSETE
ID   NXSO_PSETE     STANDARD;      PRT;    15 AA.
AC   P59073;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Short neurotoxin N2 (Alpha neurotoxin) (Fragment).
OS   Pseudonaja textilis (Eastern brown snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Acanthophiinae; Pseudonaja.
OX   NCBI_TaxID=8673;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Venom;
RX   MEDLINE=99449602; PubMed=10518793;
RA   Gong N.L., Armugam A., Jeyaseelan K.;
RT   "Postsynaptic short-chain neurotoxins from Pseudonaja textilis: cDNA
RT   cloning, expression and protein characterization.";
RL   Eur. J. Biochem. 265:982-989(1999).
CC   -!- FUNCTION: Lethal neurotoxin, binds and inhibits nicotinic
CC       acetylcholine receptors (nAChR).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- MASS SPECTROMETRY: MW=6345; METHOD=Electrospray.
CC   -!- MISCELLANEOUS: LD(50) is 0.80 mg/kg by intravenous injection.
CC   -!- SIMILARITY: Belongs to the snake toxin family.
DR   InterPro; IPR003571; Snake_toxin.
DR   PROSITE; PS00272; SNAKE_TOXIN; PARTIAL.
KW   Toxin; Neurotoxin; Postsynaptic neurotoxin;
KW   Acetylcholine receptor inhibitor; Multigene family.
FT   UNSURE         3       3
FT   UNSURE        13      13
FT   NON_TER       15      15
SQ   SEQUENCE   15 AA;  1727 MW;  E149FD4BFD1EF0DD CRC64;

  Query Match             18.4%;  Score 21;  DB 1;  Length 15;
  Best Local Similarity   75.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CYDG 5
              || |
Db          3 CYKG 6


RESULT 15
UN01_PINPS
ID   UN01_PINPS     STANDARD;      PRT;    15 AA.
AC   P81106;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Unknown protein from 2D-page of needles (S1247/S1248) (N150/N151)
DE   (Fragment).
OS   Pinus pinaster (Maritime pine).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX   NCBI_TaxID=71647;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Needle;
RA   Plomion C., Costa P., Bahrman N., Frigerio J.M.;
RT   "Genetic analysis of needle proteins in maritime pine. 1. Mapping
RT   dominant and codominant protein markers assayed on diploid tissue, in
RT   a haploid-based genetic map.";
RL   Silvae Genet. 46:161-165(1997).
RN   [2]
RP   SEQUENCE.
RC   TISSUE=Needle;
RX   MEDLINE=99274088; PubMed=10344291;
RA   Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA   Frigerio J.-M., Plomion C.;
RT   "Separation and characterization of needle and xylem maritime pine
RT   proteins.";
RL   Electrophoresis 20:1098-1108(1999).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 5.5, ITS MW IS: 62 kDa.
FT   NON_TER        1       1
FT   NON_TER       15      15
SQ   SEQUENCE   15 AA;  1670 MW;  29CE44CD51E98FCF CRC64;

  Query Match             18.4%;  Score 21;  DB 1;  Length 15;
```

Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          7 SVNNDE 12
              |  |||
Db          9 SATNDE 14
```

Search completed: December 9, 2003, 14:14:36
Job time : 13 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 14:12:40 ; Search time 20 Seconds
(without alignments)
96.169 Million cell updates/sec

Title: US-09-651-685A-5
Perfect score: 114
Sequence: 1 CCYDGASVNNDETCEQRAAR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 3883

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31.5	27.6	16	2	A59042	alpha-conotoxin Ep
2	29.5	25.9	18	1	A58589	alpha-conotoxin EI
3	28	24.6	20	2	B44907	manganese peroxida
4	26	22.8	13	1	NTKNAS	alpha-conotoxin SI
5	25	21.9	16	2	A59045	alpha-conotoxin Au
6	25	21.9	16	2	A54877	alpha-conotoxin Pn
7	25	21.9	18	1	QHEC2	heat-stable entero
8	25	21.9	20	2	PC2248	lambda 112 protein
9	24	21.1	14	2	A48389	leiurutoxin III -
10	24	21.1	16	2	PD0002	inulin fructotrans
11	23	20.2	12	2	S28215	glucan endo-1,3-be
12	23	20.2	14	1	NTKN1M	alpha-conotoxin MI
13	23	20.2	16	2	PH1317	Ig heavy chain DJ
14	23	20.2	17	2	S60171	sex-lethal protein
15	23	20.2	18	2	D49570	plasma membrane ca
16	23	20.2	18	2	B48839	T-cell receptor be
17	23	20.2	19	2	PS0236	trypsin inhibitor
18	23	20.2	20	2	PC1150	equinatoxin 1B - s
19	22	19.3	12	1	A53709	alpha-conotoxin Im
20	22	19.3	13	2	A60379	factor X activator
21	22	19.3	19	2	PH1330	Ig heavy chain DJ
22	22	19.3	20	2	S08605	hypothetical prote
23	22	19.3	20	2	H49164	chromogranin-B - r
24	22	19.3	20	2	A31516	lectin, galactose/
25	21.5	18.9	13	2	B58533	CD61 homolog - cha
26	21	18.4	10	2	PC4442	cytochrome c553 -
27	21	18.4	11	2	D45900	complement C3b rec
28	21	18.4	15	2	S49409	H+-transporting tw
29	21	18.4	15	2	PH1318	Ig heavy chain DJ
30	21	18.4	16	2	C59045	alpha-conotoxin Au
31	21	18.4	16	2	B54877	alpha-conotoxin Pn
32	21	18.4	16	2	H49039	T-cell receptor be
33	21	18.4	16	2	PL0137	protein kinase, 80
34	21	18.4	16	4	I79565	hypothetical TCL3/
35	21	18.4	17	2	I58087	ryanodine receptor
36	21	18.4	18	2	JU0125	polyphemusin II -
37	21	18.4	19	2	C21182	4K prothoracicotro
38	21	18.4	19	2	A21182	4K prothoracicotro
39	21	18.4	19	2	A49254	TcR delta chain V-
40	21	18.4	19	2	S65435	manganese peroxida
41	21	18.4	20	2	JT0410	bombyxin-IV chain
42	21	18.4	20	2	D25507	proteinase inhibit
43	21	18.4	20	2	S48746	beta tubulin b-bet
44	20	17.5	8	2	D61512	variant surface gl
45	20	17.5	11	2	S65395	chemical-sense-rel

ALIGNMENTS

RESULT 1
A59042
alpha-conotoxin EpI - cone shell (Conus episcopatus)
C;Species: Conus episcopatus (bishop's cone)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 13-Aug-1999
C;Accession: A59042
R;Loughnan, M.; Bond, T.; Atkins, A.; Cuevas, J.; Adams, D.J.; Broxton, N.M.; Livett, B.
J. Biol. Chem. 273, 15667-15674, 1998
A;Title: Alpha-conotoxin EpI, a novel sulfated peptide from Conus episcopatus that selec
A;Reference number: A59042; MUID:98288307; PMID:9624161
A;Accession: A59042
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-16 <LOU>
C;Superfamily: alpha-conotoxin
C;Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neurot
F;1-16/Product: alpha-conotoxin EpI #status experimental <MAT>
F;2-8,3-16/Disulfide bonds: #status experimental
F;15/Binding site: sulfate (Tyr) (covalent) #status experimental
F;16/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 27.6%; Score 31.5; DB 2; Length 16;
Best Local Similarity 40.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

```
Qy          1 CCYD-GASVNNDETC 14
              || |   ::|| : |
Db          2 CCSDPRCNMNNPDYC 16
```

RESULT 2
A58589
alpha-conotoxin EI - cone shell (Conus ermineus)
C;Species: Conus ermineus (ermine cone)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A58589
R;Martinez, J.S.; Olivera, B.M.; Gray, W.R.; Craig, A.G.; Groebe, D.R.; Abramson, S.N.;
Biochemistry 34, 14519-14526, 1995
A;Title: alpha-Conotoxin EI, a new nicotinic acetylcholine receptor antagonist with nove
A;Reference number: A58589; MUID:96062516; PMID:7578057
A;Accession: A58589
A;Molecule type: protein
A;Residues: 1-18 <MAR>
A;Note: sequence confirmed by chemical synthesis
C;Comment: This alpha-conotoxin, as an acetylcholine receptor inhibitor, is a postsynapt
C;Superfamily: alpha-conotoxin
C;Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; hydroxyproline; pos
F;3/Modified site: 4-hydroxyproline (Pro) #status experimental
F;4-10,5-18/Disulfide bonds: #status experimental
F;18/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 25.9%; Score 29.5; DB 1; Length 18;

```
  Best Local Similarity   33.3%;  Pred. No. 3.8e+02;
  Matches    5;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 CCY-DGASVNNDETC 14
              |||     :::| : |
Db          4 CCYHPTCNMSNPQIC 18


RESULT 3
B44907
manganese peroxidase (EC 1.11.1.13) H3 - basidiomycete (Phanerochaete chrysosporium) (fr
C;Species: Phanerochaete chrysosporium
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 06-Dec-1996
C;Accession: B44907
R;Pease, E.A.; Tien, M.
J. Bacteriol. 174, 3532-3540, 1992
A;Title: Heterogeneity and regulation of manganese peroxidases from Phanerochaete chrysc
A;Reference number: A44907; MUID:92276331; PMID:1592808
A;Accession: B44907
A;Molecule type: protein
A;Residues: 1-20 <PEA>
A;Experimental source: BKM-F-1767, ATCC 24725
A;Note: sequence extracted from NCBI backbone (NCBIP:104607)
C;Superfamily: lignin peroxidase
C;Keywords: extracellular protein; glycoprotein; heme; oxidoreductase

  Query Match             24.6%;  Score 28;  DB 2;  Length 20;
  Best Local Similarity   71.4%;  Pred. No. 7.2e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 DGASVNN 10
              ||  |||
Db          5 DGTKVNN 11


RESULT 4
NTKNAS
alpha-conotoxin SIA - cone shell (Conus striatus)
C;Species: Conus striatus (striated cone)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 23-May-1997
C;Accession: A40312
R;Myers, R.A.; Zafaralla, G.C.; Gray, W.R.; Abbott, J.; Cruz, L.J.; Olivera, B.M.
Biochemistry 30, 9370-9377, 1991
A;Title: alpha-Conotoxins, small peptide probes of nicotinic acetylcholine receptors.
A;Reference number: A40312; MUID:91369955; PMID:1892838
A;Accession: A40312
A;Molecule type: protein
A;Residues: 1-13 <MYE>
C;Comment: This paralytic toxin from a fish-hunting cone snail inhibits the acetylcholir
C;Superfamily: alpha-conotoxin
C;Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neurot
F;2-7,3-13/Disulfide bonds: #status experimental
F;13/Modified site: amidated carboxyl end (Cys) #status experimental

  Query Match             22.8%;  Score 26;  DB 1;  Length 13;
  Best Local Similarity   36.4%;  Pred. No. 9.8e+02;
  Matches    4;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNND 11
              ||:     | |
Db          2 CCHPACGKNFD 12


RESULT 5
A59045
alpha-conotoxin AuIA - cone shell (Conus aulicus)
C;Species: Conus aulicus (court cone)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999
C;Accession: A59045
R;Luo, S.; Kulak, J.M.; Cartier, G.E.; Jacobsen, R.B.; Yoshikami, D.; Olivera, B.M.; McI
J. Neurosci. 18, 8571-8579, 1998
A;Title: Alpha-conotoxin AuIB selectively blocks alpha3beta4 nicotinic acetylcholine rec
A;Reference number: A59045; MUID:99003392; PMID:9786965
A;Accession: A59045
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-16 <LUO>
C;Superfamily: alpha-conotoxin
C;Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neurot
F;1-16/Product: alpha-conotoxin AuIA #status experimental <MAT>
F;2-8,3-16/Disulfide bonds: #status experimental
F;16/Modified site: amidated carboxyl end (Cys) #status experimental

  Query Match             21.9%;  Score 25;  DB 2;  Length 16;
  Best Local Similarity   28.6%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    1;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETC 14
              | |       | : |
Db          3 CSYPPCFATNSDYC 16


RESULT 6
A54877
alpha-conotoxin PnIA [validated] - cone shell (Conus pennaceus)
N;Alternate names: alpha-CTx-PnIA
C;Species: Conus pennaceus
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 15-Sep-2000
C;Accession: A54877
R;Fainzilber, M.; Hasson, A.; Oren, R.; Burlingame, A.L.; Gordon, D.; Spira, M.E.; Zlotki
Biochemistry 33, 9523-9529, 1994
A;Title: New mollusc-specific alpha-conotoxins block Aplysia neuronal acetylcholine recep
A;Reference number: A54877; MUID:94347719; PMID:8068627
A;Accession: A54877
A;Molecule type: protein
A;Residues: 1-16 <FAI>
R;Hu, S.H.; Gehrmann, J.; Guddat, L.W.; Alewood, P.F.; Craik, D.J.; Martin, J.L.
submitted to the Brookhaven Protein Data Bank, January 1996
A;Reference number: A66355; PDB:1PEN
A;Contents: annotation; X-ray crystallography, 1.1 angstroms; residues 1-16
C;Comment: This alpha-conotoxin, as an acetylcholine receptor inhibitor, is a postsynapti
C;Superfamily: alpha-conotoxin
C;Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neurot
F;2-8,3-16/Disulfide bonds: #status experimental
F;16/Modified site: amidated carboxyl end (Cys) #status experimental

  Query Match             21.9%;  Score 25;  DB 2;  Length 16;
  Best Local Similarity   28.6%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVNNDETC 14
              |     : || : |
Db          3 CSLPPCAANNPDYC 16


RESULT 7
QHEC2
heat-stable enterotoxin ST-2 - Escherichia coli
C;Species: Escherichia coli
C;Date: 06-Jul-1982 #sequence_revision 06-Jul-1982 #text_change 31-Dec-1996
C;Accession: A01823
R;Chan, S.K.; Giannella, R.A.
J. Biol. Chem. 256, 7744-7746, 1981
A;Title: Amino acid sequence of heat-stable enterotoxin produced by Escherichia coli pat
A;Reference number: A01823; MUID:81264141; PMID:7021541
A;Accession: A01823
A;Molecule type: protein
A;Residues: 1-18 <CHA>
A;Experimental source: strain 18D, serotype 0.42:k86:H37
C;Comment: This enterotoxin is one of several, of differing molecular sizes, produced by
idues of the heat-stable enterotoxin ST-I.
C;Superfamily: heat-stable enterotoxin ST
C;Keywords: enterotoxin; heat-stable protein
F;1-18/Product: heat-stable enterotoxin ST-2 #status experimental <MAT>
F;5-10,6-14,9-17/Disulfide bonds: #status predicted
```

```
  Query Match             21.9%;  Score 25;  DB 1;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCY 3
              |||
Db          9 CCY 11


RESULT 8
PC2248
lambda 112 protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 17-Sep-1997
C;Accession: PC2248
R;Sakai, N.; Inui, K.; Fujii, N.; Fukushima, H.; Nishimoto, J.; Yanagihara, I.; Isegawa,
Biochem. Biophys. Res. Commun. 198, 485-491, 1994
A;Title: Krabbe disease: isolation and characterization of a full-length cDNA for human
A;Reference number: JC2397; MUID:94128088; PMID:8297359
A;Accession: PC2248
A;Molecule type: mRNA
A;Residues: 1-20 <SAK>
A;Cross-references: DDBJ:D25284

  Query Match             21.9%;  Score 25;  DB 2;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCY 3
              |||
Db          5 CCY 7


RESULT 9
A48389
leiurutoxin III - Egyptian scorpion (fragment)
C;Species: Leiurus quinquestriatus (Egyptian scorpion)
C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 23-Mar-1995
C;Accession: A48389
R;Valdivia, H.H.; Martin, B.M.; Escobar, L.; Possani, L.D.
Biochem. Int. 27, 953-962, 1992
A;Title: Noxiustoxin and leiurutoxin III, two homologous peptide toxins with binding pro
A;Reference number: A48389; MUID:93075256; PMID:1280139
A;Accession: A48389
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <VAL>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:117118)

  Query Match             21.1%;  Score 24;  DB 2;  Length 14;
  Best Local Similarity   35.7%;  Pred. No. 2.1e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    6;  Gaps    1;

Qy          2 CYDGASVNNDETCE 15
              ||| :       ||
Db          7 CYDSSQ------CE 14


RESULT 10
PD0002
inulin fructotransferase (depolymerizing, difructofuranose-1,2':2,3'-dianhydride-forming
C;Species: Bacillus sp.
C;Date: 10-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 03-Jun-2002
C;Accession: PD0002
R;Kang, S.I.; Kim, W.P.; Chang, Y.J.; Kim, S.I.
Biosci. Biotechnol. Biochem. 62, 628-631, 1998
A;Title: Purification and properties of inulin fructotransferase (DFA III-producing) fro
A;Reference number: PD0002
A;Accession: PD0002
A;Molecule type: protein
A;Residues: 1-16 <KAN>
C;Keywords: glycosyltransferase; hexosyltransferase

  Query Match             21.1%;  Score 24;  DB 2;  Length 16;
  Best Local Similarity   50.0%;  Pred. No. 2.4e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 DGASVNNDET 13
              ||| :|   |
Db          5 DGAPLNQVNT 14


RESULT 11
S28215
glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39) GII - barley (fragment)
N;Alternate names: (1-3)-beta-D-glucanase GII
C;Species: Hordeum vulgare (barley)
C;Date: 19-Mar-1997 #sequence_revision 17-Jul-1998 #text_change 07-May-1999
C;Accession: S28215
R;Hrmova, M.; Fincher, G.B.
Biochem. J. 289, 453-461, 1993
A;Title: Purification and properties of three (1->3)-beta-D-glucanase isoenzymes from y
A;Reference number: S28214; MUID:93143715; PMID:8424790
A;Accession: S28215
A;Molecule type: protein
A;Residues: 1-12 <HRM>
A;Experimental source: cultivar Clipper
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

  Query Match             20.2%;  Score 23;  DB 2;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 2.6e+03;
  Matches    4;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 CYDGASVN 9
              || |   |
Db          4 CYGGIGNN 11


RESULT 12
NTKN1M
alpha-conotoxin MI - cone shell (Conus magus)
C;Species: Conus magus (magus cone)
C;Date: 18-Apr-1984 #sequence_revision 18-Apr-1984 #text_change 16-Jul-1999
C;Accession: A01784
R;McIntosh, M.; Cruz, L.J.; Hunkapiller, M.W.; Gray, W.R.; Olivera, B.M.
Arch. Biochem. Biophys. 218, 329-334, 1982
A;Title: Isolation and structure of a peptide toxin from the marine snail Conus magus.
A;Reference number: A90071; MUID:83073458; PMID:7149738
A;Accession: A01784
A;Molecule type: protein
A;Residues: 1-14 <MCI>
R;Gray, W.R.; Rivier, J.E.; Galyean, R.; Cruz, L.J.; Olivera, B.M.
J. Biol. Chem. 258, 12247-12251, 1983
A;Title: Conotoxin MI. Disulfide bonding and conformational states.
A;Reference number: A92396; MUID:84032400; PMID:6630187
A;Contents: annotation; disulfide bonds
C;Comment: This alpha-conotoxin, as an acetylcholine receptor inhibitor, is a postsynap
C;Superfamily: alpha-conotoxin
C;Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neuro
F;3-8,4-14/Disulfide bonds: #status experimental
F;14/Modified site: amidated carboxyl end (Cys) #status experimental

  Query Match             20.2%;  Score 23;  DB 1;  Length 14;
  Best Local Similarity   33.3%;  Pred. No. 3e+03;
  Matches    3;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CCYDGASVN 9
              ||:     |
Db          3 CCHPACGKN 11


RESULT 13
PH1317
Ig heavy chain DJ region (clone C527-112) - human (fragment)
C;Species: Homo sapiens (man)
```

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PH1317
R;Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A;Title: Predominance of fetal type DJH joining in young children with B precursor lymph
A;Reference number: PH1302; MUID:93094761; PMID:1460419
A;Accession: PH1317
A;Molecule type: DNA
A;Residues: 1-16 <WAS>
C;Keywords: heterotetramer; immunoglobulin

Query Match 20.2%; Score 23; DB 2; Length 16;
Best Local Similarity 50.0%; Pred. No. 3.4e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 CYDGAS 7
              || |::
Db          5 CYGGST 10
```

RESULT 14
S60171
sex-lethal protein - fruit fly (Drosophila melanogaster) (fragment)
C;Species: Drosophila melanogaster
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 17-Mar-1999
C;Accession: S60171
R;Hoshijima, K.; Kohyama, A.; Watakabe, I.; Inoue, K.; Sakamoto, H.; Shimura, Y.
Nucleic Acids Res. 23, 3441-3448, 1995
A;Title: Transcriptional regulation of the Sex-lethal gene by helix-loop-helix proteins.
A;Reference number: S60171; MUID:96032836; PMID:7567454
A;Accession: S60171
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-17 <HOS>
A;Cross-references: EMBL:D50435
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995
C;Genetics:
A;Gene: FlyBase:Sxl
A;Cross-references: FlyBase:FBgn0003659

Query Match 20.2%; Score 23; DB 2; Length 17;
Best Local Similarity 66.7%; Pred. No. 3.6e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy         13 TCEQRA 18
              || ||:
Db         12 TCVQRS 17
```

RESULT 15
D49570
plasma membrane calcium pump - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 12-Dec-1997
C;Accession: D49570
R;Stauffer, T.P.; Hilfiker, H.; Carafoli, E.; Strehler, E.E.
J. Biol. Chem. 268, 25993-26003, 1993
A;Title: Quantitative analysis of alternative splicing options of human plasma membrane
A;Reference number: A49570; MUID:94064681; PMID:8245032
A;Accession: D49570
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-18 <STA>
A;Experimental source: cerebral cortex
A;Note: sequence inconsistent with nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:139934, NCBIP:139939)
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C;Keywords: ATP

Query Match 20.2%; Score 23; DB 2; Length 18;
Best Local Similarity 75.0%; Pred. No. 3.8e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 CYDG 5
              |:||
Db          5 CWDG 8
```

Search completed: December 9, 2003, 14:15:55
Job time : 21 secs